EXHIBIT 1

REDACTED

ATTORNEYS' EYES ONLY

IN THE UNITED STATES DISTRICT COURT
FOR THE CENTRAL DISTRICT OF CALIFORNIA

BERNADINE GRIFFITH, et al., individually and on behalf of all others similarly situated, Plaintiffs, v. TIKTOK, INC., a corporation; BYTEDANCE, INC., a corporation, Defendants.	Case No. 5:23-cv-00964-SB-E

EXPERT REBUTTAL REPORT OF ZUBAIR SHAFIQ, PH.D.

October 11, 2024

Table of Contents

Table of Figures

Table of Appendices

Appendix	Description
A	CV of Zubair Shafiq, Ph. D. (served on September 20, 2024)
B	**10_11 Supplemented** Script for Extracting Unmatched Data in Produced March 28 and May 21 Processed Data
C	Pixel Analysis on Random Sample of Websites (served on September 20, 2024)
D	**10_11 Supplemented** [redacted] in Produced March 28 and May 21 Processed Data
E	**10_11 Supplemented** Unique Website Domains in Produced March 28 and May 21 Processed Unmatched Data
F	**10_11 Supplemented** Unique [redacted] in Produced March 28 and May 21 Processed Unmatched Data
G	**10_11 Supplemented** Unique [redacted] in Produced March 28 and May 21 Processed Unmatched Data
H	Categorization of Plaintiff Internet Artifact Data (served on September 20, 2024)
I	**10_11 Supplemented** Unique [redacted] in Produced March 28 and May 21 Processed Unmatched Data
J	**10_11 Supplemented** Unique [redacted] in Produced March 28 and May 21 Processed Unmatched Data
K	**10_11 Supplemented** Script for Data Containing [redacted] in Produced March 28 and May 21 Processed Unmatched Data
L	**10_11 Supplemented** Average and Standard Deviation of Events Per 3p Cookie in Produced March 28 and May 21 Processed Unmatched Data
M	**10_11 Supplemented** Script for Extracting [redacted] in Produced March 28 and May 21 Processed Unmatched Data
N	Privacy Policies of Top Websites in Produced March 28 and May 21 Processed Unmatched Pixel Data (served on September 20, 2024)

O	URL Classification of 10,000 Random URL Samples from Produced March 28 and May 21 Processed Unmatched Pixel Data (served on September 20, 2024)
P	[redacted] in Produced March 28 and May 21 Processed Unmatched Pixel Data (served on September 20, 2024)
Q	Script for Extracting Unique URLs in Produced March 28 and May 21 Processed Unmatched Pixel Data (served on September 20, 2024)
R	Supplemented URLs in [redacted] March 28 and May 21 Processed Unmatched Pixel Data (served on September 20, 2024 and supplemented on September 30, 2024)
S	"Vignette" Analysis of TikTok's Collection of Data on One Non-TikTok User on March 28, 2024 and May 21, 2024 (served on September 20, 2024)

I. QUALIFICATIONS AND ASSIGNMENT

1. My name is Zubair Shafiq, Ph.D. I am an Associate Professor of Computer Science at the University of California-Davis. I have been retained by counsel for Plaintiffs to serve as an independent expert in this litigation. On June 21, 2024, I submitted the Declaration of Zubair Shafiq, Ph.D., in Support of Plaintiffs' Motion for Class Certification ("Opening Class Certification Declaration"). On July 26, 2024, I submitted the Reply Declaration of Zubair Shafiq, Ph.D., in Support of Plaintiffs' Motion for Class Certification ("Reply Class Certification Declaration"). On September 20, 2024, I submitted the Expert Report of Zubair Shafiq, Ph.D. ("Opening Report"). On September 30, 2024, I submitted a Declaration in Support of Plaintiffs' Opposition to Summary Judgment ("Summary Judgment Declaration"). I incorporate by reference the previous submissions into this Rebuttal Report, except where I have made adjustments to certain findings in the Opening Report and Summary Judgment Declaration, as I have described in this Rebuttal Report.

2. The Opening Report contained a section outlining my expertise, and my detailed CV was included as Appendix A to that report. As before, I am compensated at the rate of $750/hour. My compensation is not dependent on and in no way affects the substance of my opinions. Nor does my compensation depend on the outcome of this proceeding. I understand that, should there be any recovery in this case, I will be excluded from any disbursement of funds.

3. I understand that on September 20, 2024, TikTok served the Declaration of Ron Schnell in Support of Defendants' Motion for Summary Judgment ("Schnell Summary Judgment Report"). I also understand that on October 5, 2024, TikTok served the Declaration of Ron Schnell in Support of Defendants' Opposition to Plaintiffs' Motion for Reconsideration of the Order Denying Motion for Class Certification ("Schnell Reconsideration Report").

4. I have been asked by counsel for Plaintiffs to review the Schnell Summary Judgment Report and the Schnell Reconsideration Report and to render any opinions that I have concerning it. My analysis and conclusions with respect to that review are set out in this Rebuttal Report. After submitting my Opening Report on September 20, I, with a team of consultants working under my supervision, have continued to analyze the two-day sample data and Plaintiffs' browsing history of websites that use or used the TikTok Pixel or TikTok Events API as of March 28 and May 21. I report my findings on this continued analysis in this Rebuttal Report.

5. In addition to the items identified in the Opening Report, a list of materials reviewed and relied upon for this rebuttal report are identified in the attached **Exhibit A**.

6. I reserve the right to amend, modify and supplement this Report and the Opening Report should new or additional information be made available to me.

II. SUMMARY OF REBUTTAL OPINIONS PROFFERED

7. In the Schnell Summary Judgment Report, submitted on September 20, 2024, Mr. Schnell concluded that [redacted] He did not bother to do so although he would have had plenty of time to review that data given the fact that the data is in the possession of TikTok. Now confronted with that evidence of [redacted], which was undertaken in an accelerated time period given TikTok's significant delay in producing that data and obstacles presented in

[1] Schnell Summary Judgment Report at Section VI header (above ¶ 27).

allowing Plaintiffs' access to that data in an efficient manner.[2] Mr. Schnell's criticisms, presented in the Schnell Reconsideration Report, however, reach only marginal issues—[redacted]—and does not and cannot challenge my core findings: [redacted]

8. Rebuttal Opinion No. 1: Mr. Schnell's argument that TikTok does not collect identifying information is meritless, and his criticisms of my analysis does not change my underlying finding that TikTok collects identifying information on non-TikTok users. In particular, Mr. Schnell's opinions about the purported privacy protections offered by hashing of plaintext email addresses and phone numbers is wrong and contradicted by the warnings of the Federal Trade Commission and peer-reviewed computer science research. Mr. Schnell also misconstrues my entropy and identifiability analysis by failing to consider the second half of that analysis: that identity resolution services match identifiers like IP address, user agent, and cookies to persons or households.

9. Rebuttal Opinion No. 2: Based on further investigation, including the fact that TikTok now uses User Agent Client Hints to collect "high entropy" device information from web browsers, I have adjusted my findings [redacted]. Even with these adjustments, it is my opinion that [redacted]

[2] *See* Shafiq Opening Report at Section IX.

[redacted]

[redacted].

10. Rebuttal Opinion No. 3: I disagree with Mr. Schnell's criticisms about the IAB classification taxonomy and my use of that taxonomy, and my opinion remains unchanged that TikTok collects sensitive data on non-TikTok users. My opinion is supported by peer-reviewed research in the field of computer science, as well as TikTok's own internal documents that reflect that the very data that Mr. Schnell argues is not private is considered by TikTok to be "private" and "protected" data—"even if de-identified, anonymized, or aggregated"—in its regular course of business.

11. Rebuttal Opinion No. 4: Even accounting for Mr. Schnell's criticisms about the script I wrote [redacted] a URL need not contain a search term to be sensitive and can indeed be sensitive even without a search term embedded in it.

12. Rebuttal Opinion No. 5: Even accounting for Mr. Schnell's criticisms about the [redacted]

[redacted]

13. Rebuttal Opinion No. 6: The Schnell Summary Judgment Report's assertions that "advertisers control the configurations of their websites and what data their web pages disclose via the TikTok Pixel or EAPI"[3] ignore fundamental aspects of how TikTok has designed the tools to work, ignore TikTok's own internal documentation about data collection via Pixel that it did not disclose to advertisers, and ignore the unrebutted default aspects of both tools. I incorporate by reference my analysis and opinions presented in my Reply Class Certification Declaration (Dkt. 198-2) and in particular paragraphs 24-36 of that Reply Declaration.

14. Rebuttal Opinion No. 7: The Schnell Summary Judgment Report's assertions that there is "no indication that the contents of Plaintiffs' communications were (nor could they be) intercepted by TikTok"[4] is based on a faulty understanding of the Pixel and a narrow definition of "communication" that is not supported in the field of computer science. In light of the evidence that the Pixel collects information even ***before*** a webpage finishes loading, as well as the commonly accepted understanding of "communication" in computer science as a webpage load session rather than a packet, it is my opinion that TikTok intercepts communications between website visitors' browsers and websites.

15. Rebuttal Opinion No. 8: I have supplemented my findings on TikTok's retention and use of non-TikTok user data based on documents recently made available for my review.

[3] Schnell Summary Judgment Report at ¶ 17.

[4] Schnell Summary Judgment Report at ¶ 19.

III. REBUTTAL TO THE SCHNELL RECONSIDERATION REPORT

16. I first offer my opinions on the Schnell Reconsideration Report. To the extent that there is overlap between the opinions that Mr. Schnell offers in his Reconsideration Report and Summary Judgment Report, I respond to those opinions in this section.

A. TikTok's Collection of Identifying and Identifiable Information

a. TikTok's Collection of Hashed PII

17. Mr. Schnell's opinions about the security of hashing[5] is wrong.

18. First, Mr. Schnell is wrong that "hashes are a form of one-way encryption."[6] He does not cite anything for this faulty understanding. In fact, it is common knowledge in the field of computer science that hashing is ***not*** encryption.[7] Unless encryption is added on top of hashing, hashing does not provide confidentiality. To elaborate, encryption uses a secret key without which the encrypted information cannot be reversed. By contrast, there is no secret key involved in the hashing of plaintext email addresses or phone numbers as done by TikTok. As I explained in my Opening Report and [redacted] hashing can be easily reversed.

19. In fact, with a median capability GPU, calculating all possible hashes to reverse around 10 billion emails or phone numbers would take only a few ***seconds***. As demonstrated in my Opening Report, phone numbers can be enumerated, and there are approximately 6 billion possible US phone numbers based on the North American Numbering Plan (NANP). Likewise,

[5] *See* Schnell Reconsideration Report at ¶¶ 17, 24-31.

[6] Schnell Reconsideration Report at ¶ 27.

[7] https://www.clickssl.net/blog/difference-between-hashing-vs-encryption ("Hashing emphasizes the integrity of the information while Encryption focuses on the confidentiality of the data."); https://www.pingidentity.com/en/resources/blog/post/encryption-vs-hashing-vs-salting.html (explaining that encryption converts data into a secure format that only those with the decryption key are allowed to access and that data is protected both in transit and at rest, while hashing transform data into a fixed-size string of characters, typically used for verifying the integrity of data and securely storing passwords).

email addresses can be enumerated based on common naming patterns or sourced from leaked email address databases.[8,9] The median GPU can compute approximately 1.8 billion hashes ***per second,***[10] which means that it would take only a few seconds to compute hashes of all emails or phone numbers. Mr. Schnell exaggerates the resources necessary to reverse hashed email addresses and phone numbers.

20. Mr. Schnell also has no response to the fact that the Federal Trade Commission (FTC) agrees with my opinions regarding hashing and its lack of privacy protection. In my Opening Report, I cited the FTC's warnings for over a decade, and most recently, that "**hashes aren't 'anonymous' and can still be used to identify users, and their misuse can lead to harm. Companies should not act or claim as if hashing personal information renders it anonymized.**"[11] In fact, according to the FTC, claiming that hashing renders PII anonymous is a deceptive claim:

> Hashing has a nice potential benefit: a hash by itself cannot easily be used to guess what the original data was. For this reason, companies often use hashing in cases where they are uncomfortable writing down or sharing the directly identifying data, but they still want to be able to store the data for matching against later. Since the hash "2813448ce6316cb70b38fa29c8c64130" appears meaningless and seemingly can't be used to find the original phone number, companies often claim that hashing allows them to preserve user privacy.
>
> **This logic is as old as it is flawed – hashes aren't "anonymous" and can still be used to identify users, and their misuse can lead to harm. Companies should not act or claim as if hashing personal information renders it anonymized. FTC staff will remain vigilant to ensure companies are following the law and take action when the privacy claims they make are deceptive.**[12]

---

[8] *See, e.g.*, https://haveibeenpwned.com/ (14 billion leaked email addresses).

[9] *See, e.g.*, https://breachdirectory.org/.

[10] *See, e.g.*, https://openbenchmarking.org/test/pts/hashcat-1.1.1 (the SHA hash rate of median GPU is 1.8 billion hashes per second).

[11] https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2024/07/no-hashing-still-doesnt-make-your-data-anonymous (quoted in Opening Report at ¶ 57).

[12] https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2024/07/no-hashing-still-doesnt-make-your-data-anonymous (emphasis in original).

21. The scientific community also agrees with my opinion that hashing does not offer privacy protection. Hashes of predictable or known input, such as email addresses or phone numbers, can be trivially reversed. This is called a "rainbow table" approach. A rainbow table is essentially a large pre-built collection of hashes for known email addresses and phone numbers. Given a hash, one can quickly search the rainbow table for a match. If they find a match, they know the original input (e.g., the email address or phone number) without having to guess it or do extra work. This means that plain hashes can be easily uncovered using these pre-computed tables of hashes. Unless additional security measures like encryption or salting are used (which TikTok does not use), hashing provides no protection.[13]

22. Even Mr. Schnell admits that services like those that I described in my Opening Report can reverse hashes of consumers' email addresses and phone numbers. His argument boils down to the fact that the services are not free[14] and that they will not be able to reverse 100% of hashes—only 70 to 90% of them.[15]

23. Finally, in a footnote, Mr. Schnell misleadingly states that "anyone who visits any website is giving that website their IP address and User Agent, just based on the way the Internet works. This is whether a pixel is in use or not."[16] To be very clear, when a Pixel is not in use, a website visitor is giving ***just that website***—the intended first-party recipient of that information as

[13] *See, e.g.*, Heinrich et al., "Efficiently Recovering Hashed Phone Numbers Leaked via Apple AirDrop," https://encrypto.de/papers/HHSSW21Demo.pdf ("we leverage a custom rainbow table construction to reverse SHA-256 hashes of phone numbers in a matter of milliseconds"); Hagen et al., "All the Numbers are US: Large-scale Abuses of Contact Discovery in Mobile Messengers," https://eprint.iacr.org/2020/1119.pdf ("Furthermore, we show that currently deployed hashing-based contact discovery protocols are severely broken by comparing three methods for efficient hash reversal of mobile phone numbers.").

[14] Schnell Reconsideration Report at ¶ 32.

[15] Schnell Reconsideration Report at ¶ 29 (Datafinder, 70%), ¶ 30 (Datazapp, 71%), ¶ 31 (The Data Group, up to 90%).

[16] Schnell Reconsideration Report at ¶ 39 n.24.

evidenced by their navigation to that website—their IP address and User Agent; the website visitor is ***not*** giving the IP address and User Agent to undisclosed third parties like TikTok.

b. *Entropy and Identity Resolution Services*

24. Mr. Schnell next argues that the "use of the word 'identifiability' or ('identifiable') as it is used in the article Dr. Shafiq cites is used in a different way than he is using it. . . . [A]fter a certain amount of entropy bits, it is likely that you could isolate that two visits by the same user are just that: the same user. It does not mean that you can tell who that user is, unless that is something you already know (such as if that user is already a TikTok account holder)."[17] Mr. Schnell ignores the second of the two-step identifiability process I describe in my Opening Report. Indeed, he addresses only Paragraphs 73-84 of my opening report while remaining unaddressed Paragraphs 85-88, which discussed how "identity resolution services match the identifiers such as IP address, user agent, and cookies to persons or households."[18] In other words, he erects another strawman and disagrees with the proposition (which I never made) that every person can be identified by his or her IP address alone.

25. The following is the opinion that I actually put forth in my Opening Report and that went unaddressed by Mr. Schnell. Entropy is a "privacy metric to quantify the risk of identifiability" and is used to assess whether a certain set of input information (IP Address, User Agent, Cookies) is sufficiently identifiable.[19] Schnell cannot and does not dispute the fact that IP Address, User Agent, and Cookies are identifiers given that the amount of entropy in them is far greater than 32 bits.[20] Given the undisputed fact that IP Address, User Agent, and Cookies are

[17] Schnell Reconsideration Report at ¶ 33.

[18] Shafiq Opening Report at ¶ 85.

[19] *See generally* Shafiq Opening Report at ¶¶ 73-84.

[20] *See, e.g.*, Schnell Reconsideration Report at ¶ 34 ("Dr. Shafiq states (correctly) that an IP address is at a minimum 32 bits.").

identifiers, one can use identity resolution services or data brokers to link the identifiers to persons or households.[21]

26. As I quoted in my Opening Report, the Interactive Advertising Bureau ("IAB") explains in its Identity Solutions Guidance that: "An ID solution is a product or a service that can help identify a person and/or household across digital environments . . . or other devices with which consumers interact and consume media."[22] IAB lists the following as "commonly used components to build and identity [sic]":

- IP Address
- MAC Address
- Hashed Email Address
- Telephone number
- User agent string
- First-party cookie
- Third-party cookie
- Link decoration[23]

27. My Opening Report further identified concrete identity resolution services like Experian, Adobe, The Trade Desk, Lotame, Xandr, and Criteo that can link identifiers like the above to identities.[24] Mr. Schnell does not address any of these portions of my Opening Report.

28. Mr. Schnell's decision to ignore the availability of identity resolution services and data brokers is all the more telling because [redacted]

[21] *See* Shafiq Opening Report at ¶ 88.

[22] https://iabtechlab.com/wp-content/uploads/2024/05/Identity-Solutions-Guidance-FINAL.pdf at 14 (quoted in Shafiq Opening Report at ¶ 86).

[23] https://iabtechlab.com/wp-content/uploads/2024/05/Identity-Solutions-Guidance-FINAL.pdf at 14-15.

[24] *See* Shafiq Opening Report at ¶ 88 & Figures 13-18.

[25] TIKTOK-BG-000002069; *see also* TIKTOK-BG-000001162 ([redacted]); TIKTOK-BG-000009921 ([redacted]; TIKTOK-BG-000643356 at -360 ([redacted].

[redacted]

29. [redacted]

c. *Prevalence of the Use of VPN*

30. Mr. Schnell states that "at least 31% of all Internet users use VPNs."[28] This statement is misleading because while some Internet users may use a VPN at some point in their life, the source nowhere states that these Internet users always or exclusively use VPN. Even if a

[26] TIKTOK-BG-000171763 at -767.

[27] TIKTOK-BG-000171763 at -771.

[28] Schnell Reconsideration Report at ¶ 37 (citing https://surfshark.com/blog/vpn-users).

third of Internet users sometimes use VPN, research shows that overall VPN usage remains a tiny fraction of all Internet traffic.[29,30]

d. [redacted] Mr. Schnell's observations and analysis overall ***reinforce*** my opinion that IP Address and User Agent are identifying.

32. Google Chrome began implementing a user-agent reduction last year, which means that it no longer provides the specific device model and operating system version information that is customarily transmitted in the User Agent field.[31] Rather, it is replacing that data that used to be found in User Agent with a "default value[,] [which] will always be Android 10 on a model K."[32] As an example, the User Agent field for someone using Android Version 14.0.0 on a Samsung Galaxy S22 Ultra (SM-S908U) device, and another person using Android Version 12.0.0 on a Samsung Galaxy S10 (SM-G973U), will both be ["Android 10; K"].

33. This does not mean, however, [redacted]

[29] Schumann, Luca, Trinh Viet Doan, Tanya Shreedhar, Ricky Mok, and Vaibhav Bajpai. Impact of evolving protocols and COVID-19 on internet traffic shares. arXiv:2201.00142 (2022).

[30] Dutkowska-Zuk, Agnieszka, Austin Hounsel, Amy Morrill, Andre Xiong, Marshini Chetty, and Nick Feamster. "How and why people use virtual private networks." In 31st USENIX Security Symposium (USENIX Security 22), pp. 3451-3465. 2022.

[31] https://www.chromium.org/updates/ua-reduction/.

[32] https://developers.google.com/privacy-sandbox/blog/user-agent-reduction-android-model-and-version.

[redacted] I have confirmed by reviewing TikTok Pixel's publicly available source code that TikTok does indeed uses User Agent Client Hints[34] to collect "high entropy" device information.[35] [redacted]

34. [redacted]

35. [redacted]

[33] *See, e.g.*, TIKTOK-BG-000158056 at -057; *see also id.* at -059 ("[redacted]

[34] https://developer.mozilla.org/en-US/docs/Web/API/User-Agent_Client_Hints_API ("The User-Agent Client Hints API extends Client Hints to provide a way of exposing browser and platform information via User-Agent response and request headers, and a JavaScript API.").

[35] https://analytics.tiktok.com/i18n/pixel/static/main.MTdkNGE4ZTU0MQ.js (navigator.userAgentData.getHighEntropyValues(["model","platformVersion"]).

[redacted]

[36] Tr. of Simran Sahni 30(b)(6) Deposition (June 18, 2024) at 173:16-18 [redacted]").

[37] Schnell Reconsideration Report at ¶ 33.

[38] [redacted]

[39] [redacted]

[40] [redacted] TIKTOK-BG-000008030 at -037 (emphasis added).

[redacted]

[41] *See, e.g.*, Reply Class Certification Report (Dkt. 198-2) at 90-91.

[redacted]

[42] *See* Griffith's Supplemental Responses to Defendants' First Set of Interrogatories (Nov. 29, 2023) at Response to ROG 11 (p.15).

[43] 10_11 Supplemented Appendix D_[redacted] in Produced Processed Data.

[redacted]

[44] 10_11 Supplemented Appendix D_[redacted] in Produced Processed Data.

[45] [redacted]. *See* 10_11 Supplemented Appendix D_[redacted] in Produced Processed Unmatched "Griffith" tab.

[redacted]

[46] *See, e.g.*, TIKTOK-BG-000002930 [redacted] TIKTOK-BG-000008030 at -037 [redacted] TIKTOK-BG-000002977 at -981 ("[redacted]"); TIKTOK-BG-000736841 at -843 ("[redacted]"); TIKTOK-BG-000009897 [redacted] TIKTOK-BG-000178768 [redacted].

[47] Schnell Reconsideration Report at ¶ 51.

[48] Shafiq Opening Report at n.166.

[redacted]

[49] *See, e.g.*, Melicher et al., "(Do Not) Track Me Sometimes: Users' Contextual Preferences for Web Tracking," https://petsymposium.org/popets/2016/popets-2016-0009.pdf (defining sensitive topics as "those which deal with financial services, medicine, health, file sharing, insurance or employment"); Ur et al., "Smart, useful, scary, creepy: perceptions of online behavioral advertising," https://dl.acm.org/doi/abs/10.1145/2335356.2335362; Leon et al., "What Matters to Users? Factors that Affect Users' Willingness to Share Information with Online Advertisers," https://dl.acm.org/doi/pdf/10.1145/2501604.2501611.

[50] See sources cited *supra* in n.49.

[51] TIKTOK-BG-003078836 at -836 (emphasis added). I inserted commas between the listed items for ease of comprehension.

[redacted]

52. Mr. Schnell also discusses the confidence scores that IAB assigns with its categorizations.[54] His criticisms are unfounded, misleading, and reveal his lack of understanding of the IAB's classification API.

53. First, Mr. Schnell accuses me simultaneously of (1) "removing" IAB's confidence scores from my output and "fail[ing] to consider them at all" and (2) using too low of a minimum confidence score, which according to Mr. Schnell is 10%.[55] Obviously, even in Mr. Schnell's world, both of these things could not be true—I either did not consider confidence scores at all or considered them but chose too low of a confidence score as a threshold for my output. In reality, neither accusation is true.

54. IAB's category classification is a multi-label classification.[56] In other words, one URL can be and often is classified into multiple categories, some sensitive and some not. The expressly stated purpose of my analysis was to report whether a URL collected by TikTok and

[52] TIKTOK-BG-000430922 at -922.

[53] TIKTOK-BG-000430922 at -922 (emphasis added).

[54] *See* Schnell Reconsideration Report at ¶¶ 53-64.

[55] *Compare* Schnell Reconsideration Report at ¶ 54 [redacted] "), *with id.* at ¶ 59 [redacted] ").

[56] Herrera, F., Charte, F., Rivera, A.J., Del Jesus, M.J., Herrera, F., Charte, F., Rivera, A.J. and del Jesus, M.J., 2016. Multilabel classification (pp. 17-31). Springer International Publishing.

[redacted]

[57] Schnell Reconsideration Report at ¶¶ 56-57.

[58] *See* ATTORNEYS' EYES ONLY Appendix D – [redacted] produced with Schnell Reconsideration Report on October 5, 2024.

[59] *See* Schnell Reconsideration Report at ¶ 60.

[redacted]

57. Mr. Schnell also cherry-picks four URLs that he notes were imprecisely classified by the API[62] while saying nothing about the [redacted]

[redacted]

[62] Schnell Reconsideration Report at ¶¶ 61-64.

[redacted]

[63] Appendix O_URL Classification from my Opening Report.

[redacted]

[64] *See* Shafiq Summary Judgment Declaration at ¶ 48.

[65] Schnell Reconsideration Report at ¶¶ 65, 72.

[redacted]

[66] Schnell Reconsideration Report at ¶ 68.

[67] [redacted]

[68] 10_11 Supplemented Appendix M_Search Terms.

[69] [redacted]

[70] [redacted]

[redacted]

[71] 10_11 Supplemented Appendix M_Search Terms. [redacted]

[72] [redacted]

[73] [redacted].

[redacted]

[redacted]

[74] Schnell Reconsideration Report at ¶¶ 75-78.

[75] Schnell Reconsideration Report at ¶¶ 79-83.

[76] This criteria was chosen based on Google's policies for its Gmail addresses (https://support.google.com/mail/answer/9211434), [redacted]

[redacted]

[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

[redacted]

[77] *See* Shafiq Opening Report at ¶ 69(a) [redacted]

[78] 10_11 Supplemented Appendix [redacted].

70. [redacted]

	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

[79] Schnell Reconsideration Report at ¶¶ 106-112.

[80] 10_11 Supplemented Appendix [redacted].

[81] 10_11 Supplemented Appendix [redacted]

[82] 10_11 Supplemented Appendix [redacted]

71. [redacted]

	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]

[redacted]

[83] 10_11 Supplemented Appendix [redacted].

[84] 10_11 Supplemented Appendix [redacted]

[85] [redacted]

b. [redacted]

86 86 [redacted]

[88] Schnell Reconsideration Report at ¶¶ 85-86.

[redacted]

[89] Schnell Reconsideration Report at ¶ 87.

[90] Schnell Reconsideration Report at ¶¶ 88-89.

[91] In my modified plaintext phone number analysis, I have included only those area codes that appear in the North American Numbering Plan (https://nationalnanpa.com/enas/geoAreaCodeNumberReport.do). This list excludes toll free numbers, unused area codes, and other area codes that may be reserved for special purposes. I have also excluded phone numbers that begin with 0 or 1 (https://www.nationalnanpa.com/about_us/abt_nanp.html).

[92] 10_11 Supplemented Appendix J_[redacted]

[redacted]

	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

[redacted]

[93] 10_11 Supplemented Appendix [redacted]

[94] Schnell Reconsideration Report at ¶ 105.

[redacted]

[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

[95] 10_11 Supplemented Appendix [redacted].

[96] 10_11 Supplemented Appendix [redacted].

E. Timing of Data Collection by Events API

78. Mr. Schnell disagrees that "the TikTok Events API is real-time and contemporaneous with the loading of the webpage."[97] In doing so, he misleadingly quotes only a portion of TikTok's own document, stating that "TikTok recommends sending the event . . . as soon as it is seen on the advertiser's server."[98] The full quote, which I understand TikTok did not even dispute at summary judgment, is as follows: Defendants enable and "highly recommend[]" to advertisers that use Events API to "***send the event in real-time*** (without batching) as soon as it is seen on the advertiser's server."[99] Mr. Schnell omitted the key language that TikTok itself highly recommends that advertisers send event data "***in real-time***." Mr. Schnell's assertion that the Events API does not work in real-time[100] is contradicted by TikTok's public documentation. Either Mr. Schnell is wrong or TikTok's own public documentation is wrong. My analysis shows that Mr. Schnell is wrong.

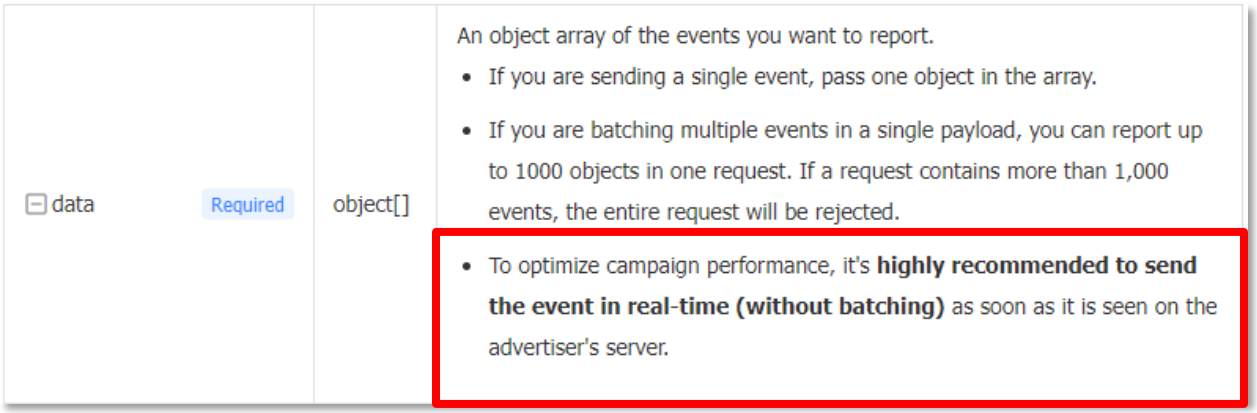


Fig. 2:
https://business-api.tiktok.com/portal/docs?rid=p41a33fdhon&id=1771100865818625

[97] Schnell Reconsideration Report at ¶ 113.

[98] Schnell Reconsideration Report at ¶ 115.

[99] *See* Dkt. 268-3 (Joint Appendix of Facts) at Fact 41.

[100] *See* Schnell Reconsideration Report at ¶ 116 [redacted]).

79. Finally, Mr. Schnell's statement that "data that is sent to TikTok from the advertisers' websites" cannot "be sent ***before*** the website itself receives it" is a red herring.[101] I never contended that Events API intercepts data ***before*** the website itself even received it.

F. Probability Calculations and Statistical Extrapolations

80. Mr. Schnell criticizes the various probability calculations I make throughout the Opening Report.[102] Before I address Mr. Schnell's specific criticisms, I address on a more general level what statistical inferences can properly be drawn from the two-day sample and what inferences cannot properly be drawn.

a. Statistical Extrapolations

81. TikTok produced two one-day time samples of non-TikTok user data—what I have been calling "two-day sample"—that I analyzed for the purposes of my Report to draw inferences about non-TikTok users at large. For example, I used the two one-day time samples to derive population-level statistics such as the percentage of non-TikTok user data collected by TikTok that contains [redacted].

82. It is appropriate to draw population-level statistics from this time sample of all non-TikTok users who were active during those two days. In fact, these statistics cover what TikTok has represented as ***all*** data from the two days, so they represent the complete data for that time period, not merely a sample. If we assume that these two days are representative of the relevant time period, then these population-level statistics are statistically also representative of the non-TikTok user population over the relevant time period. Since TikTok did not preserve and produce all non-TikTok user data over the relevant time period, Plaintiffs have no choice but to assume

[101] *See* Schnell Reconsideration Report at ¶ 114 (emphasis added).

[102] Schnell Reconsideration Report at ¶¶ 23, 117-129.

that the data from these two days is representative of the relevant time period for the purpose of drawing ***population-level*** statistics.

83. By contrast, when analyzing this two one-day time sample of all non-TikTok user data, it is noteworthy that the absence of a ***particular*** non-TikTok user, e.g., one of the Plaintiffs, in this time sample does not necessarily mean that TikTok did not collect data from that Plaintiff. In plain terms, the absence of evidence is not evidence of absence. We can compute population-level statistics from this this two one-day time sample of all non-TikTok user data, but we cannot draw conclusions about specific non-TikTok users. Just because TikTok did not collect data from a particular non-TikTok user during those two days does not mean that TikTok did not collect data from that non-TikTok user over the relevant time period spanning multiple years. More generally, attempting to draw any statistical inferences about a particular non-TikTok user from the two one-day time samples would be unsound, an example of the "ecological fallacy."[103] (Of course, if the two-day sample happens to contain actual data from a plaintiff, then it follows that TikTok collected data from the plaintiff. That is not a statistical inference, but a direct observation.)

84. To illustrate this point, consider a hypothetical researcher studying compliance with traffic laws. He obtains a one-week sample of driving data on drivers in the United States. Drivers in the sample ran, on average, 0.23 stop signs during that week. Assuming the sample is representative (and setting aside various technical statistical issues), the researcher could infer that the average US driver runs roughly 1 stop sign per month and 12 stop signs per year. These are statistical inferences about the general population of US drivers (with a certain confidence level, subject to a margin of error, and so on). But now suppose the researcher wants to know how often

[103] *See, e.g.*, Freedman, "Ecological Inference and the Ecological Fallacy," https://web.stanford.edu/class/ed260/freedman549.pdf.

a particular individual, Jones, runs stop signs. In the one-week sample, Jones did not run a single stop sign. But it would be erroneous to infer from that sample that Jones never runs a stop sign. Similarly, if Jones ran 1 stop sign in the one-week sample, it would be erroneous to infer that Jones runs 52 stop signs per year. The one-week sample would not support those types of inferences about Jones or other specific individuals.

85. The more relevant evidence to support the conclusion that the Plaintiffs represent the population of non-TikTok users at large comes not from the two one-day time samples from TikTok but rather from their [redacted] demonstrates the generalizability of their experiences to the larger non-TikTok user population. [redacted], there might be a concern about whether their experience could generalize to all non-TikTok users. However, the Plaintiffs' browsing histories show [redacted]

b. *Mr. Schnell's Criticisms about Probability Calculations*

86. Mr. Schnell provides no basis for his various criticisms, and it is my opinion that his say-so is unfounded. For instance, Mr. Schnell states that my calculations "assume that each user visits a completely new website every time or rotates in a predictable way between websites

[104] *See* Appendix R Supplement that was filed with the Court at Dkt. 269-38.

where the Pixel might transmit an e-mail address and websites where the Pixel will never transmit an e-mail address."[105] He also states that my calculations "represent[] an upper bound" and that "[i]n reality, each time a user revisits a safe website, they're essentially 'wasting' a visit in terms of potential transmission of their e-mail address."[106] Finally, Mr. Schnell states that my calculations "assume a predictable split between safe sites and unsafe sites, which is not realistic."[107] Mr. Schnell offers no additional explanation beyond these confusing sentences which, to borrow Mr. Schnell's own words, I "struggled to understand."[108] Unlike my report which was accompanied by scripts and data outputs that Mr. Schnell could review, however, he did not provide any such supporting material beyond his say-so.[109]

87. In any event, I did not assume that users "rotate in a predictable way" between websites or a "predictable split between safe sites and unsafe sites," whatever that means. All probabilities in my Opening Report and in this Rebuttal Report are derived ***based on TikTok's own sample data***.

G. "Provocative" Content Information

88. Mr. Schnell acknowledges that the data that TikTok collected from non-TikTok users in just two days is "provocative" but insists that this "provocative" material cannot be found in the URL itself.[110] This is incorrect on at least two levels.

---

[105] Schnell Reconsideration Report at ¶ 124.

[106] Schnell Reconsideration Report at ¶ 125.

[107] Schnell Reconsideration Report at ¶ 126.

[108] Schnell Reconsideration Report at ¶ 97.

[109] Beyond these three unsupported sentences, Mr. Schnell's Reconsideration Report is littered with unsupported assumptions about online behavior with no supporting citations whatsoever. *See, e.g.*, Schnell Reconsideration Report at ¶ 121 ("In the real world, this is demonstrably not the case, and does not consider user behavior at all."), ¶ 124 ("I know from experience (and the data I've seen) that people often revisit the same sites."), ¶ 126 ("In reality, I know from my experience (and the data I've seen) that people's browsing habits vary widely."). Mr. Schnell leaves to readers to divine what exactly is the "data he's seen."

[110] Schnell Reconsideration Report at ¶ 130.

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

90. Second, Mr. Schnell is wrong when he argues that "[y]ou would have to click on the actual link to see what is on the website" and that this "is not part of the data that are sent to TikTok."[111] In making this argument, Mr. Schnell ignores Content Information that TikTok automatically collects via Pixel. As explained in my previous declarations and Opening Report, TikTok does collect Content Information corresponding to each URL.[112] TikTok's crawlers and event listeners scrape the content of each webpage the Pixel fires on.[113] TikTok's own internal

[111] Schnell Reconsideration Report at ¶ 130.

[112] *See, e.g.*, Shafiq Opening Class Certification Declaration at ¶ 65.

[113] *See* Shafiq Opening Class Certification Declaration at ¶¶ 65, 67-71, 78.

[redacted]

IV. REBUTTAL TO THE SCHNELL SUMMARY JUDGMENT REPORT

92. After considering Mr. Schnell's Summary Judgment Report, the conclusions in my Opening Report are unchanged. I discuss the flaws in the Schnell Summary Judgment Report and the bases of my disagreement with him in more detail below.

A. The Schnell Summary Judgment Report Fails to Analyze [redacted] Before Wrongly Concluding that There Is No Evidence of Data Collection from Plaintiffs

93. [redacted]

[redacted]

[114] TIKTOK-BG-000132449.

[115] TIKTOK-BG-000132448.

[116] TIKTOK-BG-000149005 at -006.

[117] TIKTOK-BG-000149005 at -006.

[118] [redacted]

[119] TIKTOK-BG-000149005 at -006.

evidence of such data-sharing before reaching that conclusion: TikTok's own data.[120] Indeed, he states that he [redacted]"[121]—with no mention of TikTok's own sample data produced in this case. While he states that "[t]hese were material provided to me by counsel for Defendants," Mr. Schnell was aware of the production of sample data that TikTok had produced in this case given the rebuttal report that he submitted in response to my Opening Class Certification Declaration. It is unclear why he forewent the opportunity to investigate the most obvious source of evidence of whether websites shared Plaintiffs' private data with TikTok. Even after I submitted my Opening Report and as part of his preparation of his Reconsideration Report, which I discuss in Section III above, Mr. Schnell appears not to have looked at TikTok's sample data itself, but rather looked only at my scripts and data outputs, which of course made the billions of rows of TikTok two-day sample data more digestible and user-friendly. Now that he has done so, as discussed above, *see supra* Sec. III.A.d, [redacted]

94. As also stated in my Opening Report (and supplemented in this Report), TikTok's two-day sample data also provided me with a list of more than [redacted] that apparently used the Pixel and/or Events API at least as of March 28, 2024 and May 21, 2024

[120] Schnell Summary Judgment Report at Section VI header (above ¶ 27).

[121] Schnell Summary Judgment Report at ¶ 28.

[redacted] These URLs include ones that the Interactive Advertising Bureau (IAB) standard content taxonomy would [redacted]

96. Since submitting my Opening Report on September 20, my team and I have continued to analyze Plaintiffs' browsing histories of websites that use the Pixel or Events API, which as I stated in my Opening Report, [redacted]

[122] Shafiq Opening Report at ¶ 64. 10_11 Supplemented [redacted]

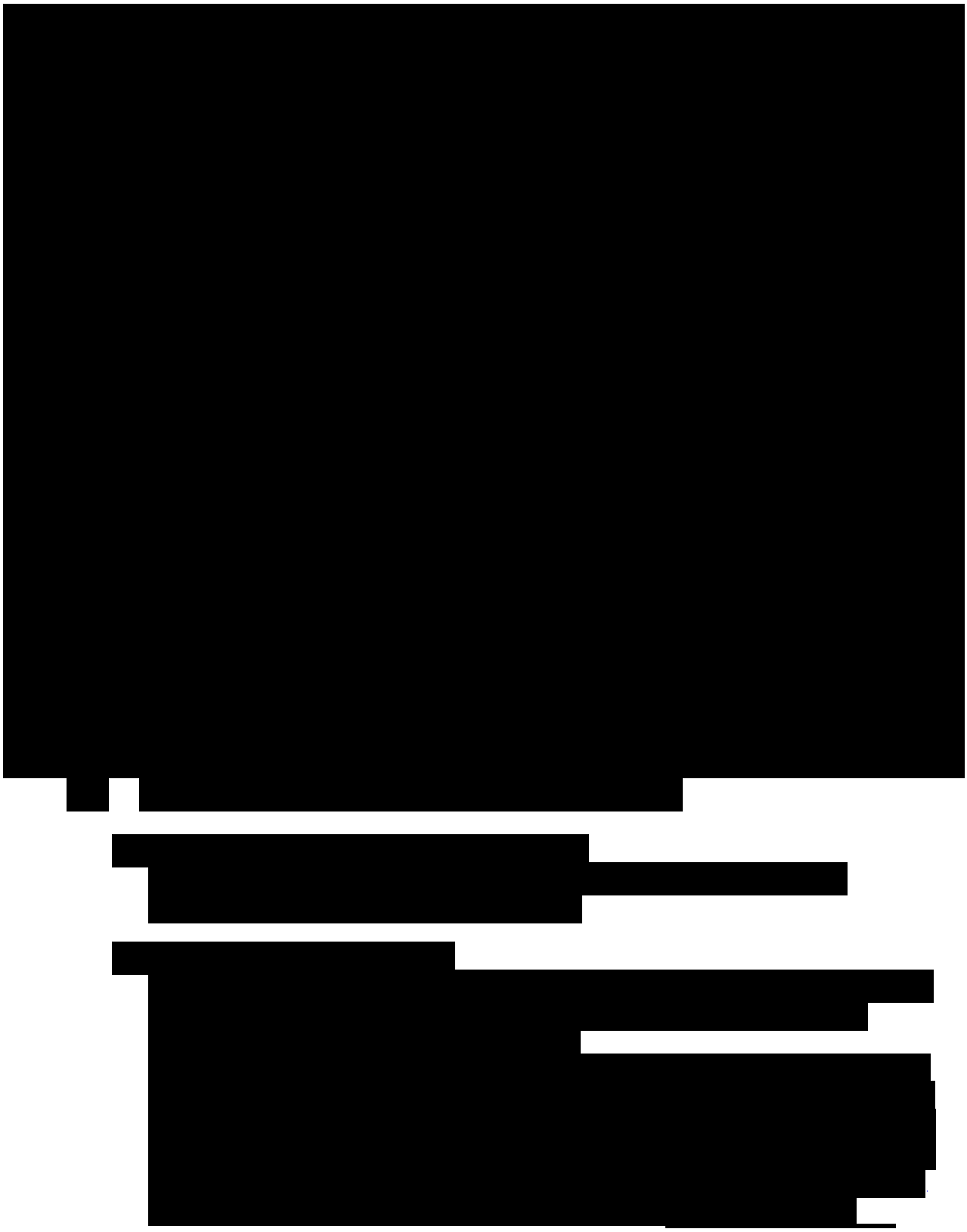
[123] These histories were attached as Appendix R to the Shafiq Opening Report and supplemented in [redacted]. It is my understanding that Appendix R Supplement was filed with the Court at Dkt. 269-38.

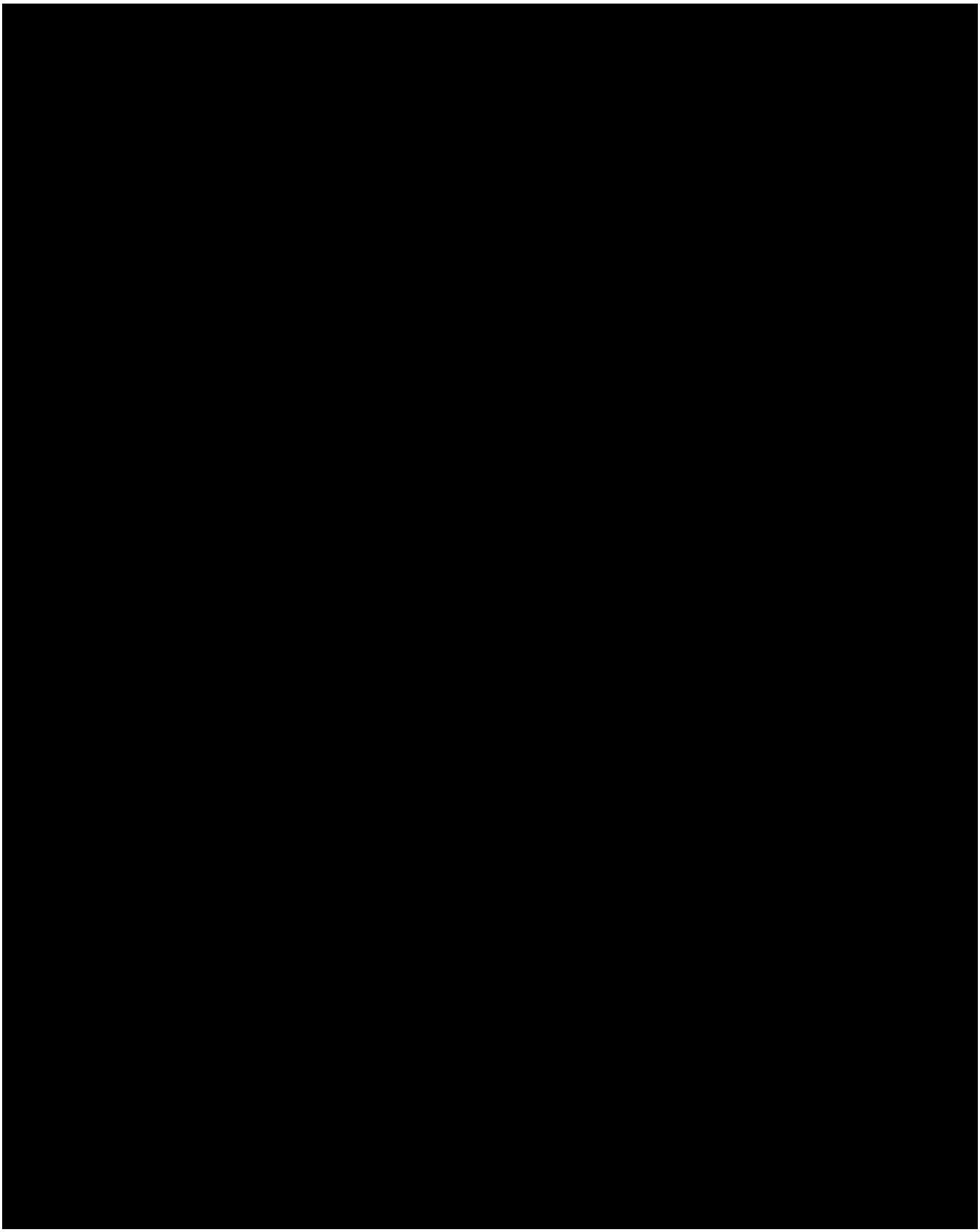
[124] 10_11 Supplemented Appendix B_[redacted]

[125] Shafiq Opening Report at ¶¶ 65-66.

[126] *See* Appendix R Supplement that was filed with the Court at Dkt. 269-38.

[redacted]



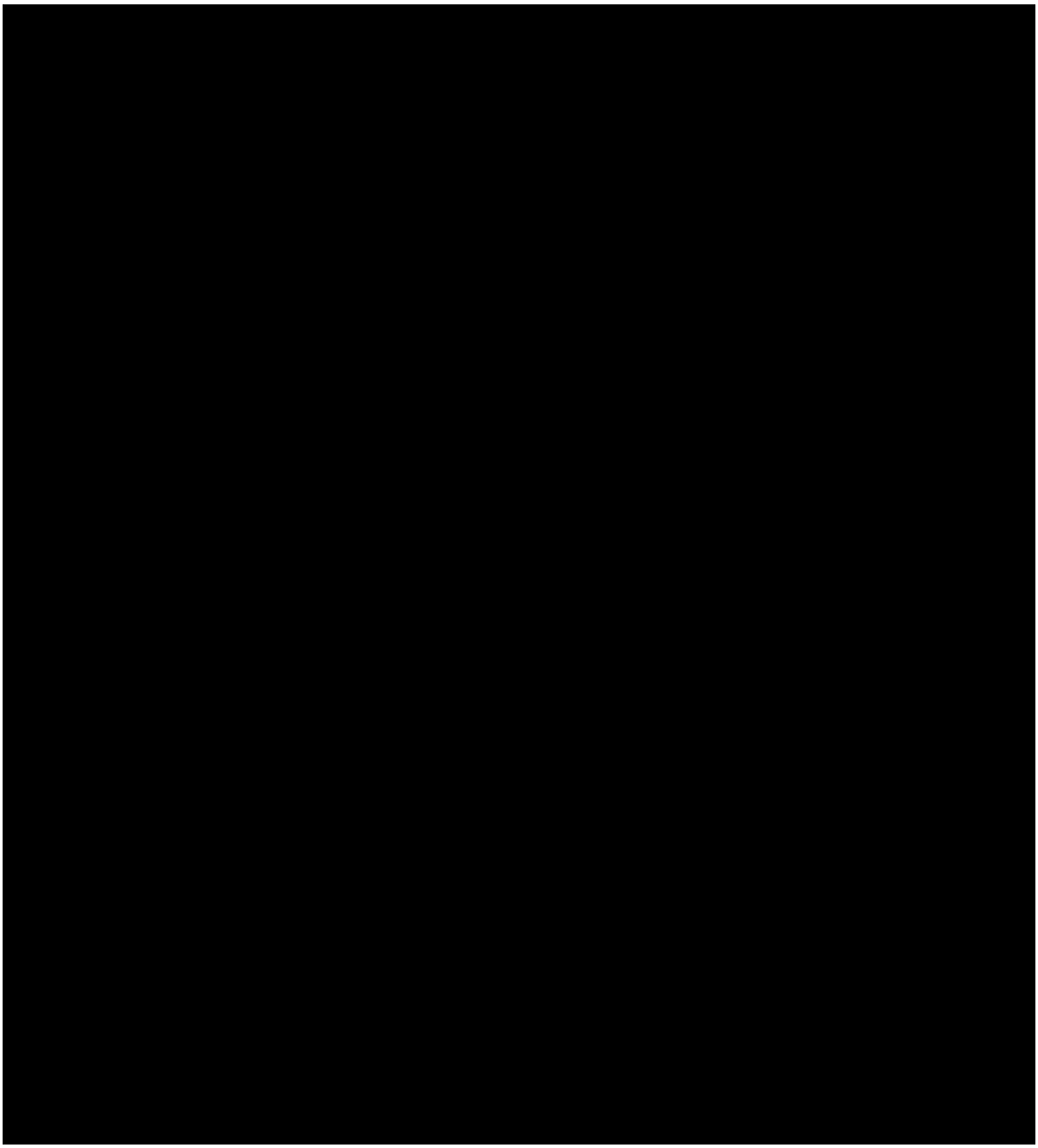


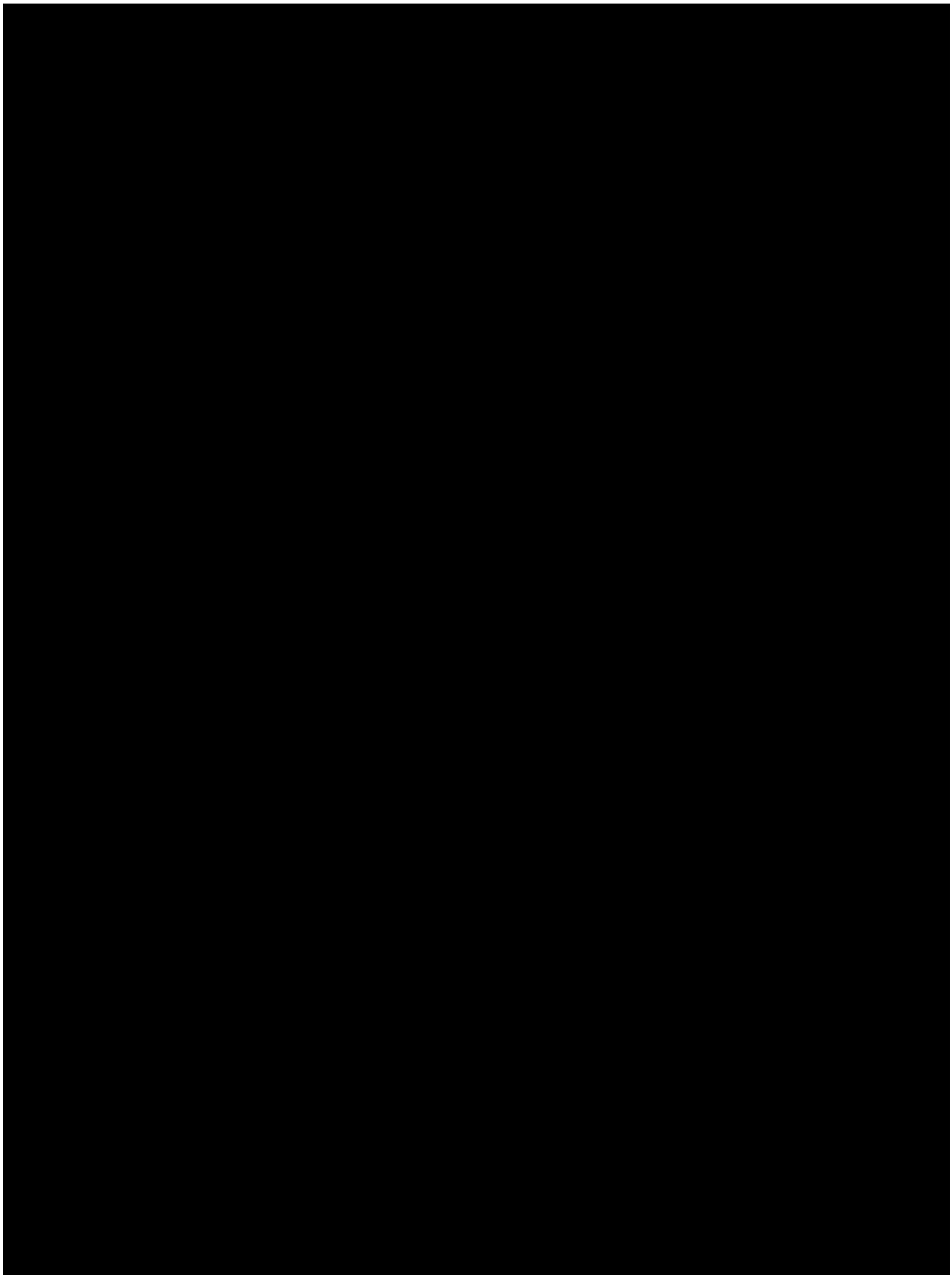
[redacted]

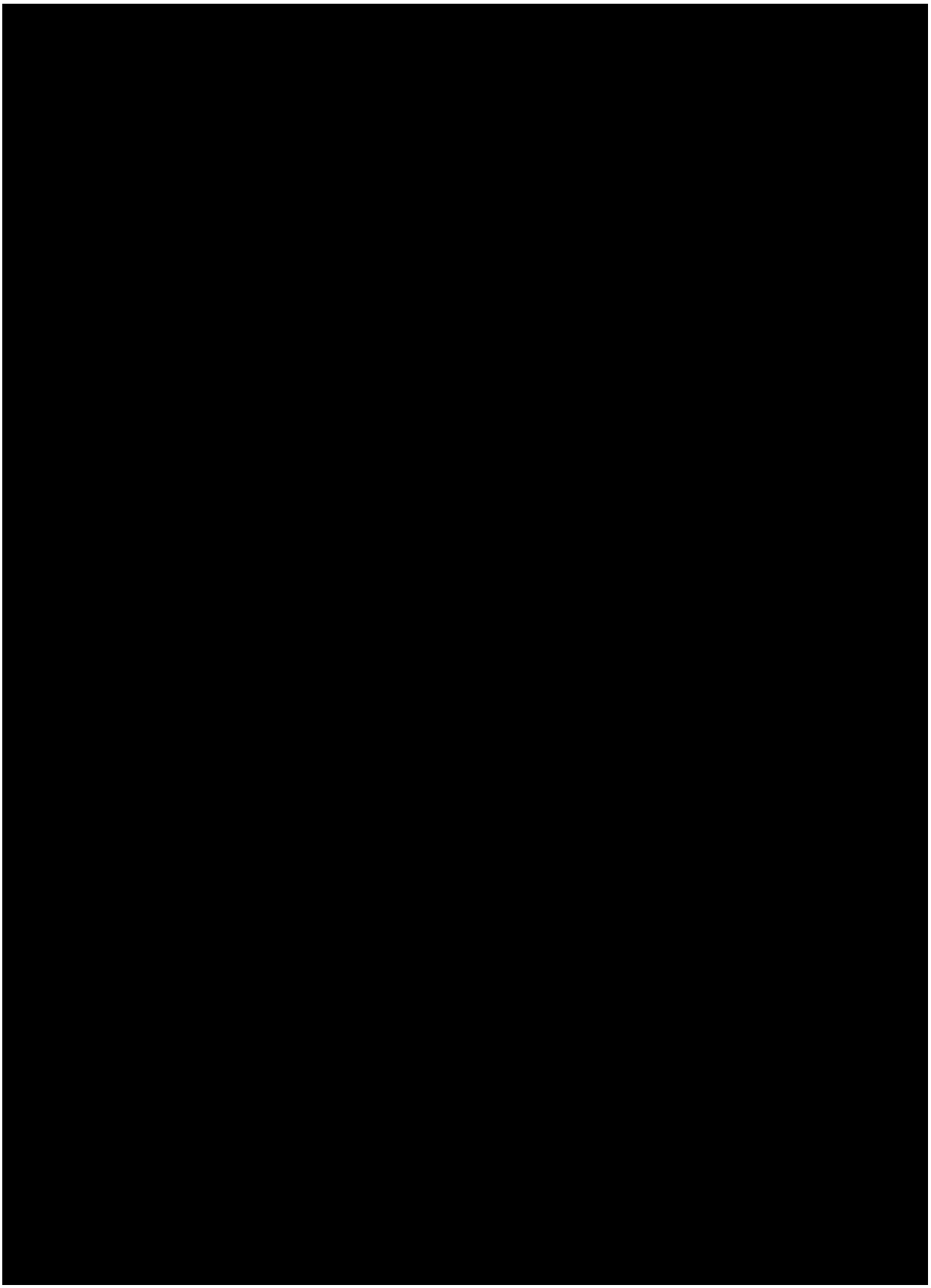
[redacted]

[127] *See* Files: "SOURCEID_07333 - [redacted] (produced as SHIH-GRIFFITHTT000183), "SOURCEID_07334 - [redacted] (SHIH-GRIFFITHTT000184), "SOURCEID_07334 - [redacted] (SHIH-GRIFFITHTT000185), "SOURCEID_07337 - [redacted] (SHIH-GRIFFITHTT000186).

[128] *See* JAE Ex. 28 at ¶ 81(a).







113. [redacted]

[redacted]

B. TikTok Controls What Data It Collects Via TikTok Pixel and Events API.

a. Advertisers Generally

114. Mr. Schnell asserts that for both Pixel and EAPI, "advertisers choose what data to share with TikTok via the Pixel and EAPI by selecting, designing, and implementing 'event' options. These events set the parameters for what data can be collected from the advertiser's potential customer."[129] This overlooks the various default types of data that Pixel collects out of the box even without the advertisers setting any events or parameters. These defaults include the [redacted] and the [redacted] all discussed in my Opening and Reply Class Certification Declarations and, as to [redacted], as recognized by Mr. Schnell himself.[130]

115. Mr. Schnell also ignores TikTok's own documents which make clear that [redacted]

116. [redacted]

[129] Schnell Summary Judgment Report at ¶ 24.

[130] Shafiq Opening Class Certification Declaration at ¶¶ 32-37 and accompanying footnotes; Reply Class Certification Declaration at ¶ 19 & nn.3-4; Schnell Class Certification Report at ¶ 41 [redacted]; Transcript of July 19, 2024 Deposition of Ron Schnell at 30:12-14 ("Q. [redacted]

[131] *See generally* Tr. of Yun-Feng Wei Deposition (July 14, 2024), at 179:18-181-20.

[redacted]

[132] TIKTOK-BG-000629566 at -566. I understand that this document was produced in Chinese, and I have been provided with a certified translation of the document.

[133] TIKTOK-BG-000629566 at -566.

[134] TIKTOK-BG-000629566 at -567, 568.

[135] TIKTOK-BG-000629566 at -567.

[136] TIKTOK-BG-003073124 at cell C3.

[redacted]

118. Mr. Schnell fails to recognize that advertisers often lack awareness regarding third-party trackers, such as TikTok Pixel, operating on their websites and their data collection practices. In a recent peer-reviewed study surveying about four hundred developers, Utz et al. reported "a widespread lack of awareness" among website developers. [138] The authors found that the website developers "either do not know or are uncertain of the true extent of data collection by the third party."[139] The authors concluded that website developers "often seem to not know or ignore the possibility that their visitors' personal data could be collected for other purposes, or simply trust the third-party service to not collect data or to employ adequate privacy protection."[140] The prevalent lack of awareness among website developers concerning data collection by third parties, such as TikTok, suggests that website developers often do not even know that their websites are using them, let alone properly configuring them and choosing what data to transmit with them.

119. Mr. Schnell also overlooks that the default data categories like timestamp, IP address, User Agent, Cookies, and certain URLs need not be collected by TikTok just because a website uses the TikTok Pixel; TikTok has deliberately and specifically implemented TikTok Pixel's source code to collect these default categories of data. In fact, Mr. Schnell fails to acknowledge that other tracking pixels do ***not*** always collect these categories of data that TikTok deliberately designed its Pixel and Events API to uniformly collect. I elaborate how each of

[137] TIKTOK-BG-003073124 at accompanying comment to cell C3.

[138] Utz, C., Amft, S., Degeling, M., Holz, T., Fahl, S., & Schaub, F. (2023). Privacy Rarely Considered: Exploring Considerations in the Adoption of Third-Party Services by Websites. Proceedings on Privacy Enhancing Technologies.

[139] *Id.*

[140] *Id.*

timestamp, IP address, User Agent, cookies, and URLs need not be collected by TikTok Pixel even where websites are configured to transmit that data to TikTok.

120. **Timestamp:** TikTok Pixel automatically includes the timestamp in the payload of HTTP POST requests from a website visitor's web browser to TikTok's server. Note that HTTP requests from a website visitor's browser to a web server do not always contain timestamp, which can be ***optionally*** included as a HTTP header.[141] For example, the HTTP requests from the website visitor's web browser to RiteAid's web server do not contain timestamp in a HTTP header. As another example, Apple's Private Click Measurement (PCM) aims to safeguard privacy "by submitting reports to the websites through an anonymization service ***after a delay***,"[142] which means that Apple would not be able to even indirectly infer exact timestamps.

121. **IP Address:** TikTok Pixel automatically includes the IP address in the network layer header of the HTTP POST request from a website visitor's web browser to TikTok's server. Note that HTTP requests from a website visitor's browser to a web server do not have to contain the IP address. To elaborate, an IP address could be "relayed" or proxied via an intermediary to mask the IP address of the original user. For example, Google's Safe Browsing service makes use of Oblivious HTTP to "hide end users' IP addresses from Google."[143] More generally, CloudFlare's Privacy Gateway allows any application or service on the Internet to make use of Oblivious HTTP to hide the IP address of website visitors. As another example, Google Analytics

[141] https://developer.mozilla.org/en-US/docs/Web/HTTP/Headers/Date ("The **Date** general HTTP header contains the date and time at which the message originated.").

[142] https://blog.mozilla.org/en/mozilla/understanding-apples-private-click-measurement/ (emphasis added).

[143] https://developers.google.com/safe-browsing/reference.

4, a tracking pixel provided by Google, also does not log or store IP addresses, nor does the iCloud Private Relay.[144]

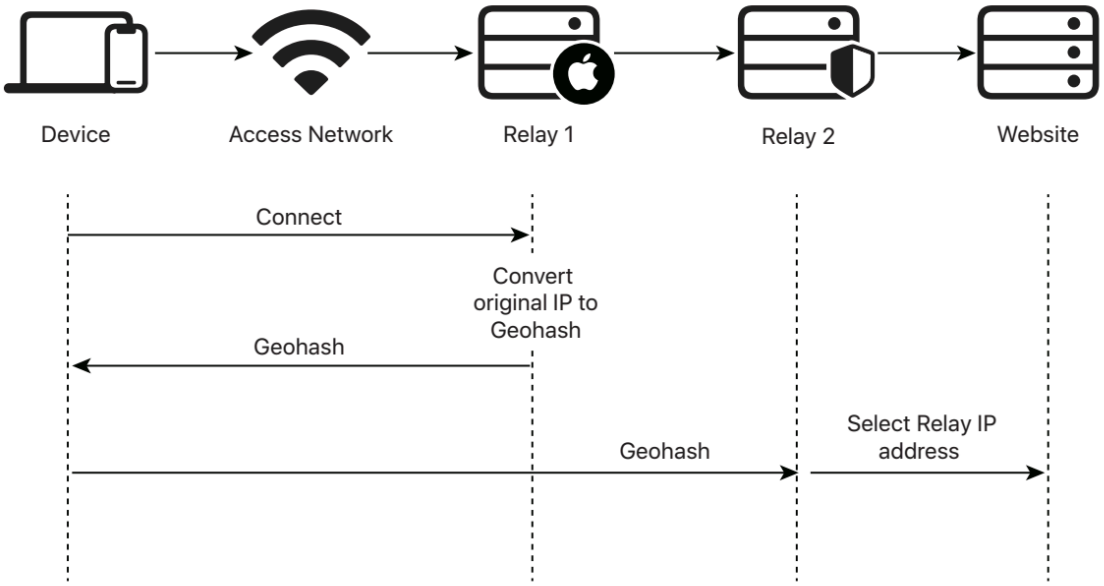


Fig. 3: How Private Relay "protect[s] users' privacy, while maintaining sufficiently accurate location information to support a personalized experience on the web"[145]

122. **User Agent:** TikTok Pixel automatically includes the user agent in the payload of HTTP POST requests from a website visitor's web browser to TikTok's server. More specifically, my review of TikTok Pixel's source code shows that TikTok Pixel purposefully collects the full user agent through a JavaScript API called Navigator.userAgent.[146] Note that HTTP requests from a website visitor's browser to a web server do not have to contain user agent. For example, the HTTP requests from the website visitor's web browser to RiteAid's web server do not contain user agent in the HTTP request payload. It is also noteworthy that the collection of user agent in the

[144] https://support.google.com/analytics/answer/12017362; https://www.apple.com/privacy/docs/iCloud_Private_Relay_Overview_Dec2021.PDF.

[145] https://www.apple.com/privacy/docs/iCloud_Private_Relay_Overview_Dec2021.PDF.

[146] https://developer.mozilla.org/en-US/docs/Web/API/Navigator/userAgent.

HTTP request payload is distinct from the user agent collected in the HTTP request header.[147] TikTok Pixel purposefully collects user agent both in the HTTP request payload and in the HTTP request header. From my experience, this is atypical.

123. **Cookies:** TikTok Pixel collects three different types of cookies: third-party cookies, first-party cookies, and session ID. Note that the HTTP requests from a website visitor's browser to a web server do not have to contain all of these cookies. More specifically, my review of TikTok Pixel's source code shows that TikTok Pixel purposefully collects the first-party cookie from the web browser's cookie jar through a JavaScript API called document.cookie[148] and sends it to TikTok's server in the "anonymous_id" field of the HTTP request payload.[149] Similarly, TikTok Pixel's source code purposefully collects the session ID from the web browser's session storage through a JavaScript API called window.sessionStorage[150] and sends it to TikTok's server in the "session_id" field of the HTTP request payload.[151]

124. **URLs:** TikTok Pixel collects the full-string URL and the referrer URL in the payload of HTTP POST requests from a website visitor's web browser to TikTok's server. More specifically, my review of TikTok Pixel's source code shows that TikTok Pixel purposefully collects the full-string URL through a JavaScript API called window.location.href[152] and sends it

[147] https://developer.mozilla.org/en-US/docs/Web/HTTP/Headers/User-Agent.

[148] https://developer.mozilla.org/en-US/docs/Web/API/Window/sessionStorage.

[149] Shafiq Opening Class Certification Declaration at ¶ 62.

[150] https://developer.mozilla.org/en-US/docs/Web/API/Document/cookie.

[151] Shafiq Opening Class Certification Declaration at ¶ 62.

[152] https://developer.mozilla.org/en-US/docs/Web/API/Window/location ("The Window.location read-only property returns a Location object with information about the current location of the document. Though Window.location is a read-only Location object, you can also assign a string to it. This means that you can work with location as if it were a string in most cases: location = 'http://www.example.com' is a synonym of location.href = 'http://www.example.com'.") TikTok Pixel source code (https://analytics.tiktok.com/i18n/pixel/static/main.MWY1ZWZmZjM0MQ.js) uses window.location to intercept the URL of the webpage.

to TikTok's server in the "url" field of the HTTP request payload.[153] Similarly, TikTok Pixel's source code purposefully collects the referrer URL through a JavaScript API called document.referrer[154] and sends it to TikTok's server in the "referrer" field of the HTTP request payload.[155] It is noteworthy that the collection of these URLs in the HTTP payload is distinct from the "Referer" collected by TikTok Pixel in the HTTP request header.[156] This distinction is particularly important because the URLs collected by TikTok Pixel through the JavaScript APIs circumvent the Referrer-Policy that can be set by website developers to limit the collection of URLs by tracking pixels.[157] As shown in the following figure, the full-string URL and the referrer URL collected by TikTok Pixel via "url" and "referrer" fields in the HTTP request payload are https://www.riteaid.com/shop/benadryl-allergy-25-mg-ultratabs-24-tablets-0033202 and https://www.riteaid.com/shop/catalogsearch/result/?q=benadryl, respectively. However, the URL collected by TikTok Pixel in the Referer header is simply https://www.riteaid.com/, which is limited due to the Referrer-Policy.[158] In other words, even where the website or browser has a Referrer-Policy in place to limit the collection of full-string URLs, TikTok Pixel nonetheless purposefully collects the full-string URLs via the "referrer" field in the HTTP request payload. This is a purposeful and intentional way in which TikTok designed its Pixel.

125. Not all pixels behave this way. There exist other tracking pixels that allow website developers to redact collection of certain information in URL. For example, Google Analytics 4

[153] Shafiq Opening Class Certification Declaration at ¶¶ 63, 68.

[154] https://developer.mozilla.org/en-US/docs/Web/API/Document/referrer ("The Document.referrer property returns the URI of the page that linked to this page.").

[155] Shafiq Opening Class Certification Declaration at ¶¶ 63, 68.

[156] https://developer.mozilla.org/en-US/docs/Web/HTTP/Headers/Referer.

[157] https://developer.mozilla.org/en-US/docs/Web/HTTP/Headers/Referrer-Policy.

[158] https://developer.mozilla.org/en-US/docs/Web/HTTP/Headers/Referrer-Policy.

provides a data-redaction feature[159] to prevent collection of specific URL query parameters. TikTok Pixel does not offer any such feature to website developers. There also exist other tracking pixels that warn website developers if they detect collection of personal data via URL.[160] TikTok Pixel does not offer any such feature to website developers.

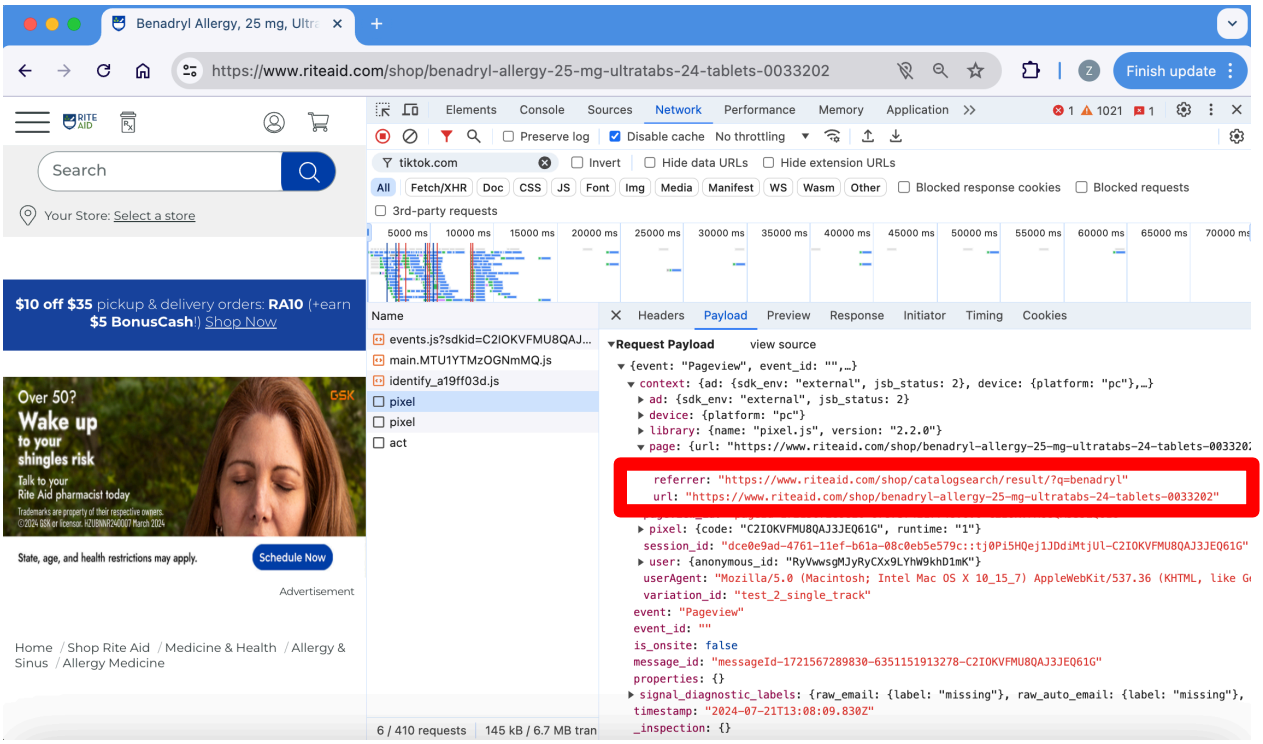


Fig. 4: Full-string URL collected by TikTok Pixel in the HTTP payload.

[159] https://support.google.com/analytics/answer/13544947.

[160] https://developers.facebook.com/community/threads/312553193608925/.

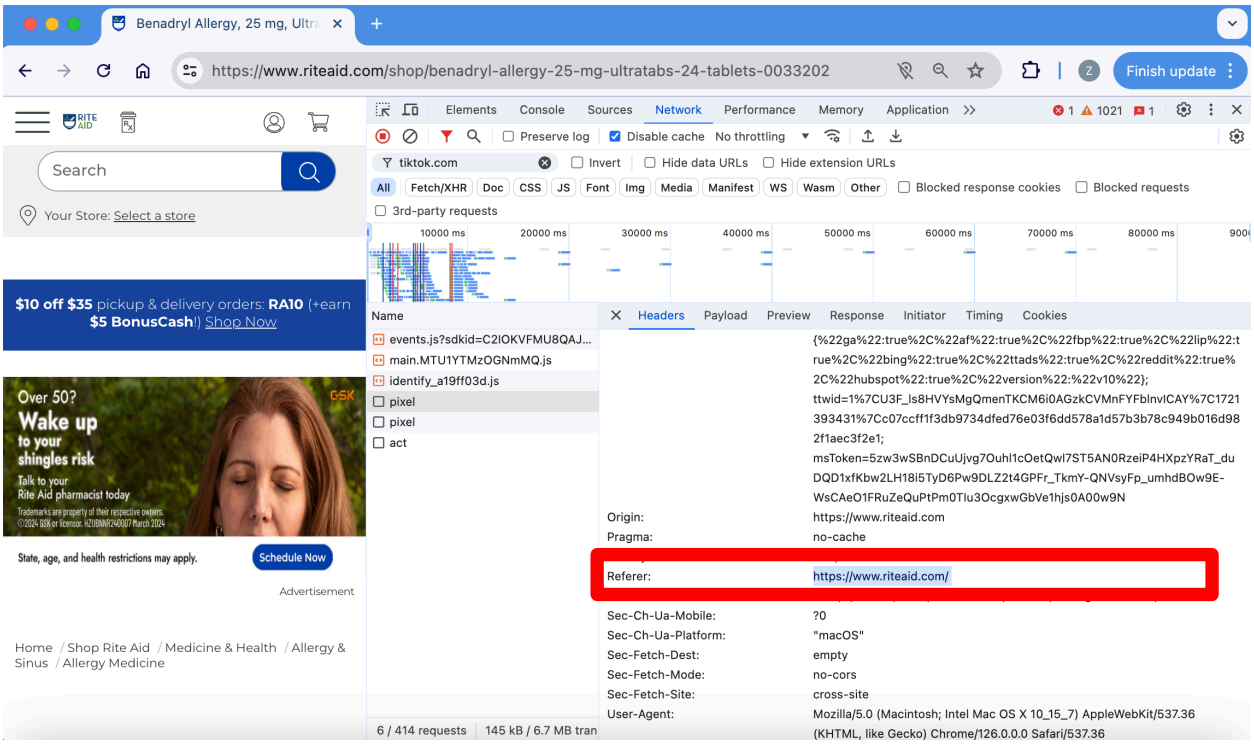


Fig. 5: The URL collected by TikTok Pixel in the HTTP Referer header (Note that it is not the full-string URL collected by TikTok Pixel in the HTTP payload.)

126. I further incorporate by reference the findings and analysis in my Reply Class Certification Declaration and, in particular, paragraphs 24-36 of that Reply Declaration.

b. *RiteAid, Hulu, Etsy, Upwork, Sweetwater, Build-a-Bear*

127. As stated in my Summary Judgment Declaration, Defendants' two-day sample data reveals that [redacted].[161] I thus disagree with Mr. Schnell's opinion that these websites "are not configured to share private data about Plaintiffs with TikTok."[162]

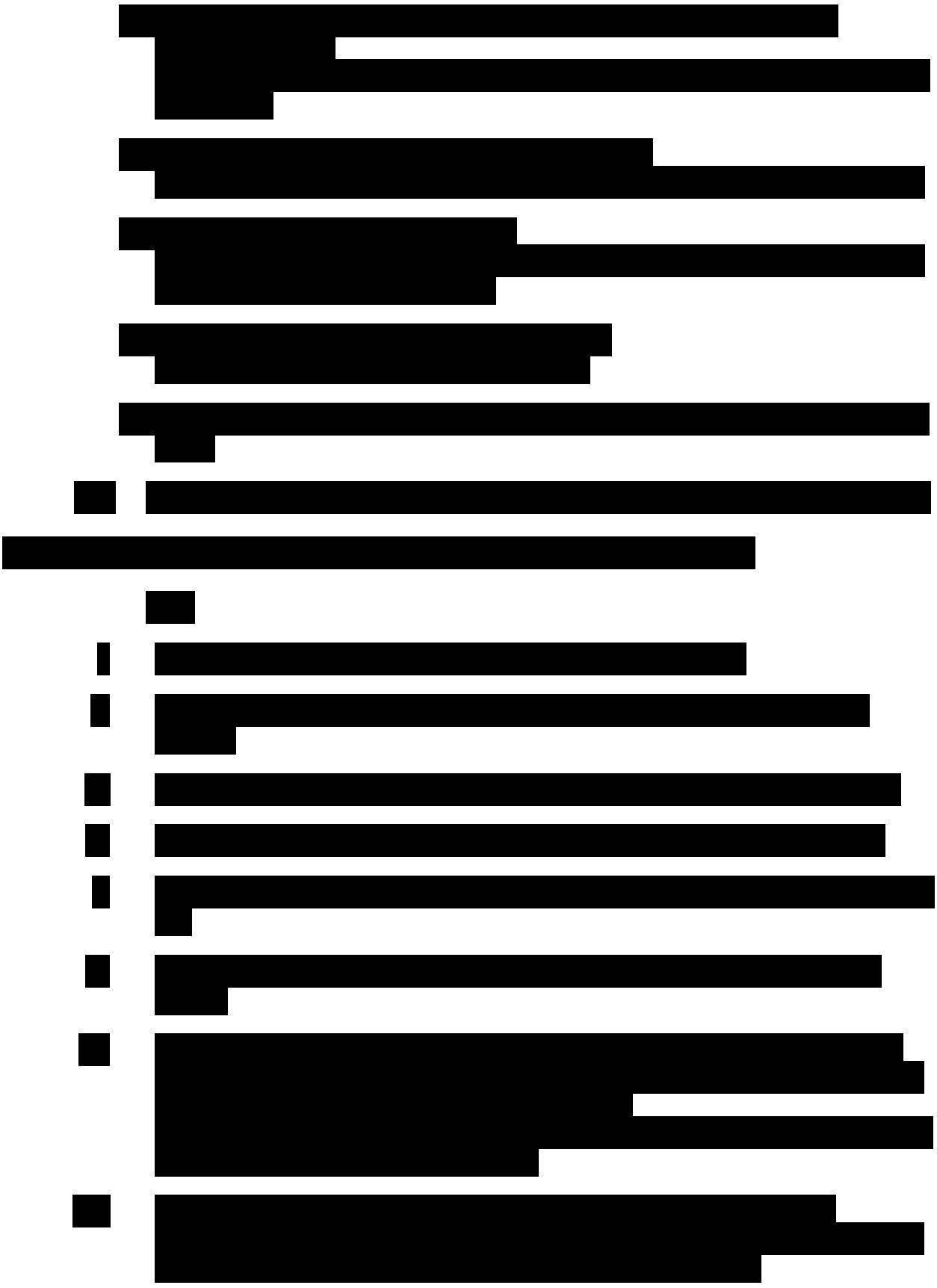
128. In the two-day sample data, I [redacted]

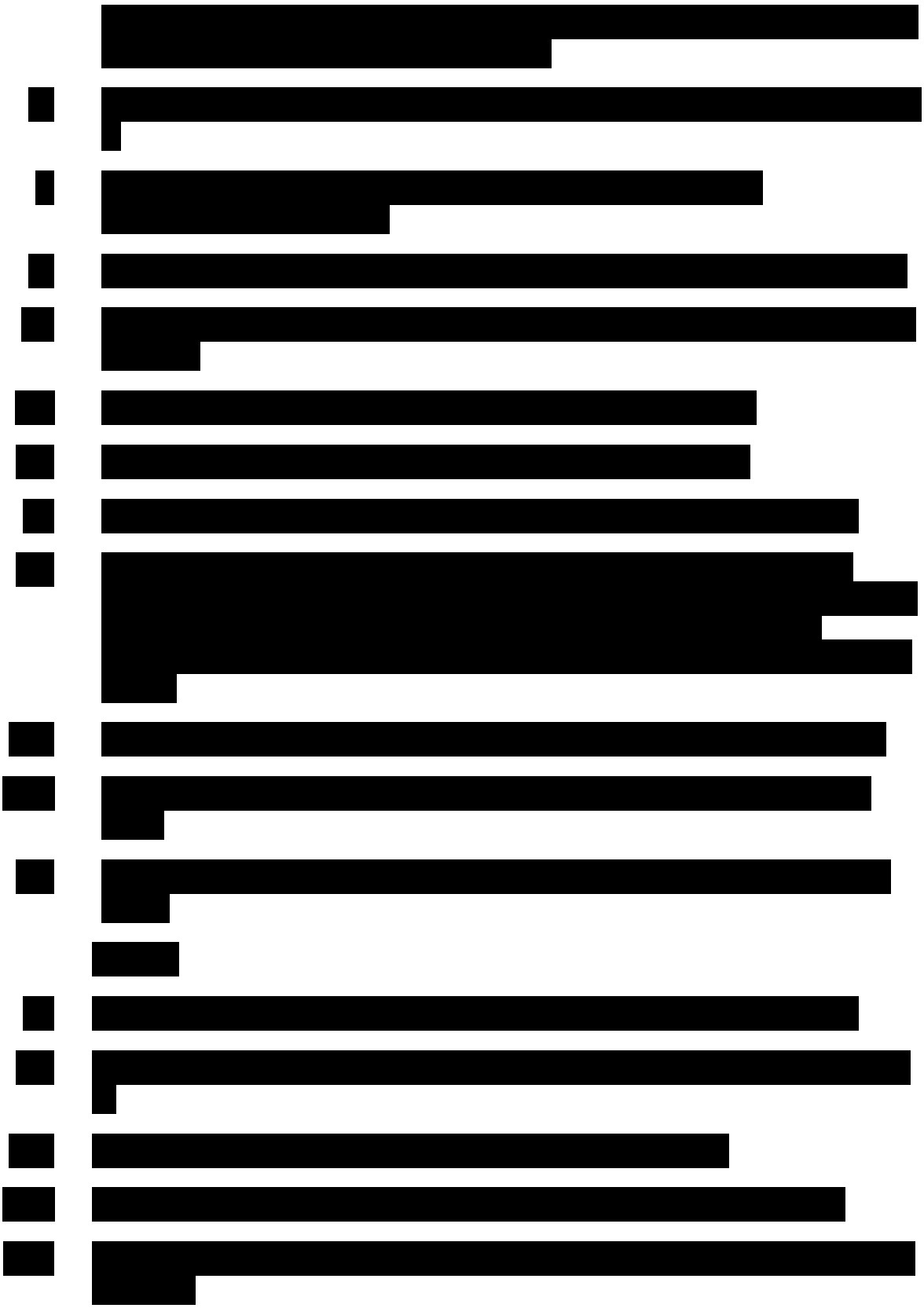
[161] *See* Shafiq Summary Judgment Declaration at ¶¶ 47-48.

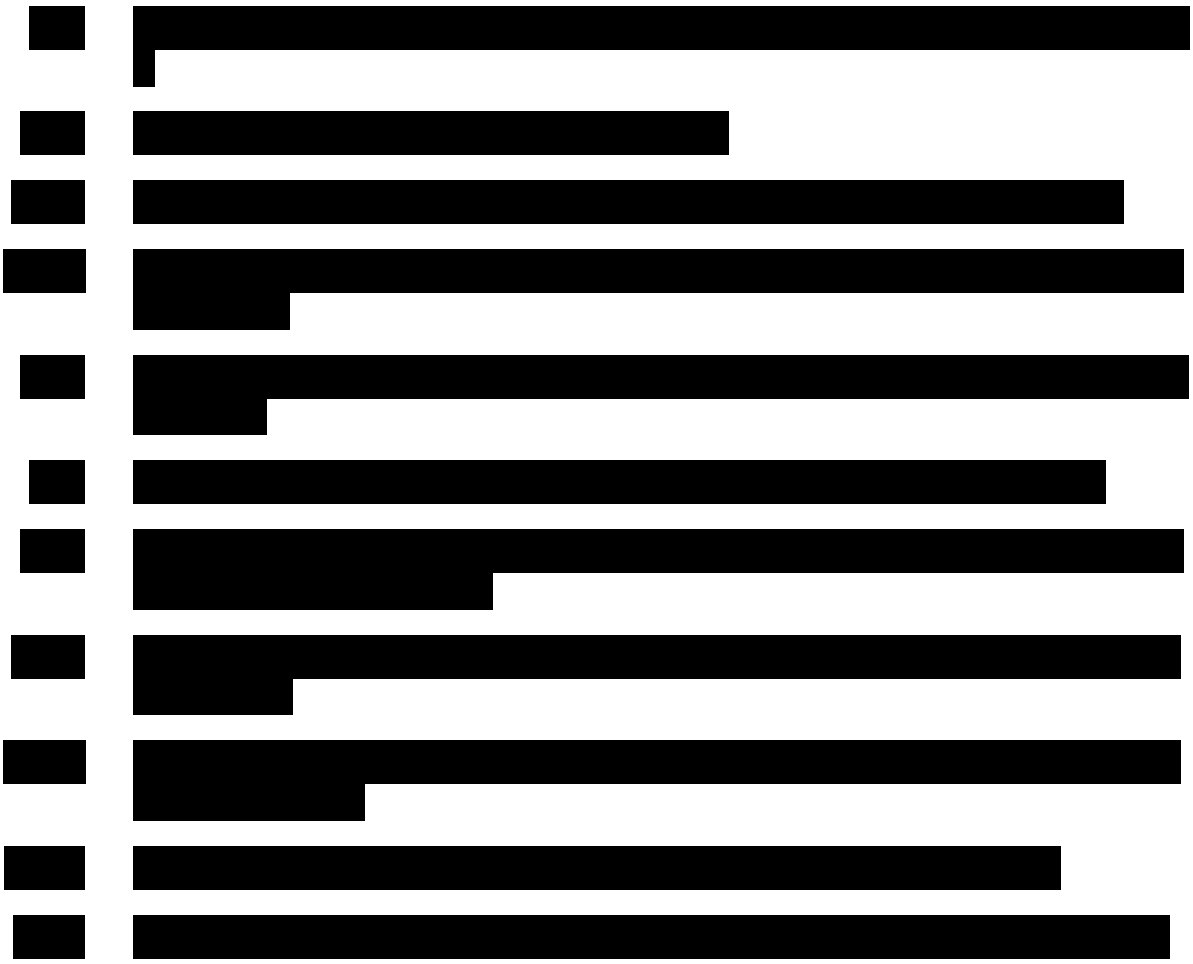
[162] Schnell Summary Judgment Declaration at Heading VI above ¶ 27.

[redacted]

[163] *See* JAE Ex. 72 (Dkt. 269-42).







C. The Schnell Summary Judgment Report Mischaracterizes Plaintiffs' Deposition Testimony

130. The report mischaracterizes Plaintiffs' deposition testimony, often without any citations to their actual deposition transcripts. I correct some of those misrepresentations below.

a. *Griffith*

131. Mr. Schnell says, without citation, that "Ms. Griffith visited Etsy's website around June 2018."[164] To the extent that Mr. Schnell suggests that Griffith visited Etsy only once, such a suggestion would be contradicted by her testimony that she visited Etsy between 20-50 times,[165]

[164] Schnell Summary Judgment Report at ¶ 51.

[165] Tr. of Bernadine Griffith Deposition (June 26, 2024) ("Griffith Tr.") at 65:15-22.

and her production showing her browsing history and purchases on Etsy since creating her Etsy account around June 2018.[166]

132. Mr. Schnell says, without citation, that Ms. Griffith "has always set cookie blockers on."[167] This ignores that Griffith's testimony was specific to ***third-party cookies***, which is a material detail given TikTok Pixel's collection of first-party cookies.[168] Further, when pressed on whether Griffith "had your cookie blocker on for every single time"[169] of the 500 times she visited Hulu, she responded "I don't know."[170]

133. Mr. Schnell also goes on, again without any citations, about how Griffith had a "0 video" watch history on Hulu even though she doesn't recall deleting her watch history as Hulu allows one to do.[171] To the extent that Mr. Schnell is suggesting that Griffith did not actually watch anything on Hulu, such a suggestion would be contradicted by her testimony that she visited Hulu approximately 500 times,[172] and that "it's my belief that I did watch them. Whether or not I actually went through the trouble to dig in and try to delete my watch history -- I believe you can. I don't recall actually doing that."[173]

---

[166] GRIFFITHTT001755- GRIFFITHTT002069; GRIFFITHTT002117.

[167] Schnell Summary Judgment Report at ¶ 90.

[168] Griffith Tr. at 61:13-19 ("Q. And you can you [sic] identify any website that you visited that uses pixel and Events API while your web or browser settings were turned on to block ***third-party cookies***? A. Again, I don't know how to separate pixels from APIs, but the websites do include Hulu, Etsy, Rite Aid, Build-A-Bear, Sweetwater." (emphasis added)).

[169] I further note that counsel's question about "cookie blocker" is an imprecise one. Based on my review of her deposition transcript, Griffith testified that she activated the third-party cookie blocker in her web browser by configuring browser settings to reject third-party cookies, as opposed to having add-on blocking software. *See* Griffth Tr. 66:15-17 ("Q. When did you put your cookie blockers in place? A. I have it set up through my web browser.").

[170] Griffith Tr. at 64:2-14.

[171] Schnell Summary Judgment Report at ¶ 90.

[172] Griffith Tr. at 63:22-64:5.

[173] Griffith Tr. at 171:16-23.

134. Finally, Mr. Schnell, citing to an inapposite portion of her deposition, misrepresents Griffith's testimony about her interaction with a cookie banner from her RiteAid visit.[174] In reality, she testified that she would click "allow" on Rite Aid's cookie banner "if that was the only choice and I could not proceed."[175] This testimony does not support Mr. Schnell's conclusion that Pixel could not collect Griffith's identifying information from her visits to Rite Aid.

b. *Watters*

135. Mr. Schnell's statement that Watters "understands that Etsy collects information, partners with third parties and shares data" is incomplete and misleading. In his deposition, Watters testified that he was ***not*** aware "explicitly what third parties" Etsy shared data with and that "if I were to have any issues with the third parties, I can't object if I don't know what the third party is."[176] Based on his review of Etsy's privacy policy presented to him by counsel at his deposition, Watters also testified that Etsy "likely" shared data with "Google and Facebook," which are disclosed by name in Etsy's privacy policy but that he "was not aware explicitly of any other third parties because they were not listed."[177]

136. [redacted] In other words, he engages in the conduct that he inaccurately accused me of engaging in, as discussed above.[178] Mr. Schnell's say-so is

[174] Schnell Summary Judgment Report at ¶ 91.

[175] Griffith Tr. at 222:15-23 ("Q. Have you ever interacted with a cookie banner from – for Rite Aid? A. I may have. I don't remember. Q. Okay. A. Probably. Q. If you did, did you click 'allow' in order to access the Rite Aid account? A. If that was the only choice and I could not proceed, then that's what I would click.").

[176] Tr. of Jacob Watters Leady Deposition (June 25, 2024) ("Watters Tr.") at 91:14-23.

[177] Watters Tr. at 92:3-10.

[178] Schnell Reconsideration Report at ¶ 51.

contradicted by peer-reviewed computer science research, which classifies data related to [redacted].[179]

c. *Shih*

137. Mr. Schnell misrepresents that Shih "uses the 'TikTok for guest' option.[180] This assertion is contradicted by both Shih's declaration and her deposition testimony. In her declaration in support of class certification, Shih stated: "I have also never watched videos on the TikTok app or website using the 'TikTok for Guest' option."[181] During her deposition, when TikTok's counsel asked Shih to define what "TikTok for guest" means, Shih responded: "It is my understanding that that is a new feature of TikTok that is -- is a way to make an account without having -- without video posting purposes. I would just say that all of the -- all of the options here basically relate to making an account. ***I have never made an account with TikTok***."[182]

138. Mr. Schnell's statement that Shih "uses 'fresh browser' to visit TikTok and the other websites at issue" is also inaccurate and misleading.[183] First of all, Mr. Schnell leaves out the time period about which Shih was testifying. TikTok's counsel asked Shih about her browsing practice "since you joined this litigation," and Shih testified that her "practice is to use a fresh browser so they can't glean where I'm coming from other than I clicked the link" in this post-litigation time period.[184] In addition, Shih testified that she uses a fresh browser when she visits

[179] *See supra* at ¶ 48; Melicher et al., "(Do Not) Track Me Sometimes: Users' Contextual Preferences for Web Tracking," https://petsymposium.org/popets/2016/popets-2016-0009.pdf (defining sensitive topics as "those which deal with financial services, medicine, health, file sharing, insurance or employment").

[180] Schnell Summary Judgment Report at ¶ 93.

[181] Dkt. 177-3 at ¶ 2.

[182] Tr. of Patricia Shih Blough Deposition (June 24, 2024) ("Shih Tr."), at 230:20-231:5 (emphasis added); *see also* Shih Tr. at 228:24-229:5 ("Q. So the activities that you have had on TikTok, that has been as a guest on TikTok? THE WITNESS: It would depend how you would define 'guest.'" (objection omitted)).

[183] Schnell Summary Judgment Report at ¶ 93.

[184] Shih Tr. at 293:10-294:8.

Etsy, Hulu, and Upwork—the websites that she was asked about during her deposition, not all websites that use or may use Pixel or Events API.

D. The Schnell Summary Judgment Report's Conclusions about Lack of Interception Are Unsupported and Contradicted by the Industry-Standard Definition of "Communication"

139. I incorporate by reference the findings and analysis in my Opening Class Certification Declaration and, in particular, paragraphs 67-71 of that Declaration. I also incorporate by reference the findings and analysis in my Reply Class Certification Declaration and, in particular, paragraphs 95-102 of that Reply Declaration.

140. Mr. Schnell is incorrect that "the Pixel can only collect information from a webpage that has loaded."[185] As I previously explained,[186] [redacted]. Thus, TikTok collects the data even if the website visitor navigates away from the webpage before the page finishes loading in the browser completely. TikTok's internal documentation confirms [redacted]

[185] Schnell Summary Judgment Report at ¶ 108.

[186] Shafiq Opening Class Certification Declaration at ¶ 70.

[187] TIKTOK-BG-000008579 at -586.

[188] TIKTOK-BG-003078303 at -308 [redacted]

[189] TIKTOK-BG-000151808 (Depo Ex. 43), at -817; *see also* Tr. of Becca Wong Depo. (May 17, 2024) at 143:9-19 [redacted]

141. Further, TikTok Pixel's source code[190] contains instructions to set up listeners for events such as "mousedown," [191] "pointerdown," [192] "mouseup," [193] "pointerup," [194] "mousemove,"[195] "pointermove,"[196] "keydown,"[197] and "click."[198] The listeners set up by TikTok Pixel are instantaneously (i.e., in real-time and without any delay) triggered when these events occur ***during the webpage load***, allowing TikTok Pixel to intercept movement of the mouse or any other pointer device, when keys are pressed, and when buttons are clicked as well as content information (e.g., using browser features such as window.location.href[199] to intercept the page

[190] https://analytics.tiktok.com/i18n/pixel/static/main.MWY1ZWZmZjM0MQ.js.

[191] https://developer.mozilla.org/en-US/docs/Web/API/Element/mousedown_event ("The mousedown event is fired at an Element when a pointing device button is pressed while the pointer is inside the element.").

[192] https://developer.mozilla.org/en-US/docs/Web/API/Pointer_events ("Pointer events are DOM events that are fired for a pointing device. They are designed to create a single DOM event model to handle pointing input devices such as a mouse, pen/stylus or touch (such as one or more fingers).").

[193] https://developer.mozilla.org/en-US/docs/Web/API/Element/mouseup_event ("The mouseup event is fired at an Element when a button on a pointing device (such as a mouse or trackpad) is released while the pointer is located inside it.").

[194] https://developer.mozilla.org/en-US/docs/Web/API/Pointer_events ("Pointer events are DOM events that are fired for a pointing device. They are designed to create a single DOM event model to handle pointing input devices such as a mouse, pen/stylus or touch (such as one or more fingers).").

[195] https://developer.mozilla.org/en-US/docs/Web/API/Element/mousemove_event ("The mousemove event is fired at an element when a pointing device (usually a mouse) is moved while the cursor's hotspot is inside it.").

[196] https://developer.mozilla.org/en-US/docs/Web/API/Pointer_events ("Pointer events are DOM events that are fired for a pointing device. They are designed to create a single DOM event model to handle pointing input devices such as a mouse, pen/stylus or touch (such as one or more fingers).").

[197] https://developer.mozilla.org/en-US/docs/Web/API/Element/keydown_event ("The keydown event is fired when a key is pressed.").

[198] https://developer.mozilla.org/en-US/docs/Web/API/Element/click_event ("An element receives a click event when any of the following occurs • a pointing-device button (such as a mouse's primary button) is both pressed and released while the pointer is located inside the element. • a touch gesture is performed on the element • the Space key or Enter key is pressed while the element is focused").

[199] https://developer.mozilla.org/en-US/docs/Web/API/Window/location ("The Window.location read-only property returns a Location object with information about the current location of the document. Though Window.location is a read-only Location object, you can also assign a string to it. This means that you can work with location as if it were a string in most cases: location = 'http://www.example.com' is a synonym of location.href = 'http://www.example.com'.") TikTok Pixel source code (https://analytics.tiktok.com/i18n/pixel/static/main.MWY1ZWZmZjM0MQ.js) uses window.location to intercept the URL of the webpage.

URL, document.referrer[200] to intercept the referrer URL, and document.querySelectorAll[201] to intercept webpage content in JSON-LD, Meta, Microdata, and OpenGraph formats) while it is in-transit from the website visitor's browser to the website's server. Thus, TikTok Pixel's event listeners assist with its interception of Content Information.

142. [redacted]

143. In addition, Mr. Schnell's assertions that "[o]n a typical website page, by the time the TikTok Pixel receives any communication from the website visitor, it is only after the website has already first received that communication, stored it, processed it, and then used it to load the next page—only then allowing the Pixel to receive that information"[203] is based on an arbitrary and narrow conception of "communication" that departs from common definition of that term in the field of computer science.

[200] https://developer.mozilla.org/en-US/docs/Web/API/Document/referrer ("The Document.referrer property returns the URI of the page that linked to this page.").

[201] https://developer.mozilla.org/en-US/docs/Web/API/Document/querySelector ("The Document method querySelector() returns the first Element within the document that matches the specified selector, or group of selectors. If no matches are found, null is returned."). TikTok Pixel source code (https://analytics.tiktok.com/i18n/pixel/static/main.MWY1ZWZmZjM0MQ.js) uses querySelectorAll to intercept content information such as JSON-LD.

[202] TIKTOK-BG-003050009 at -010 (emphasis added).

[203] Schnell Summary Judgment Report at ¶ 110.

144. Mr. Schnell appears to define the unit of "communication" as a single packet of data sent from the website visitor's web browser to the website.[204] However, in the context of web browsing, the fundamental unit of "communication" is more accurately and commonly defined as a web session, encompassing the entire sequence of interactions between the browser and the server, rather than isolating each individual HTTP "request" or "response" packet.[205,206] A web session represents the entire interaction between a website visitor's web browser and a website, beginning when the web browser sends the first HTTP request to the website's server and continuing through various HTTP requests/responses exchanged to fully load the web page. This series of HTTP request/response packets are needed to fully render the webpage in a user's web browser. Put simply, each HTTP request/response packet is a part of a web session.[207]

145. Mr. Schnell also offers no reasoned basis that he considers only data sent from website visitor to the website to be a "communication" while excluding data sent in the reverse direction—from the website to the web browser.[208] Again, in the common definition of "communication" as a session, all interactions or exchanges between the browser and website are considered "communications."[209]

---

[204] *Id.* at 108-109; *see also* Schnell Class Certification Rebuttal Report at ¶ 122 (discussing communications in terms of transmission of "packets to the server computer," "data contained in these packets" and the receipt of the packet by the host of the website).

[205] A typical HTTP session: https://http.dev/session.

[206] A typical HTTP session: https://developer.mozilla.org/en-US/docs/Web/HTTP/Session.

[207] Meiss, M., Duncan, J., Gonçalves, B., Ramasco, J.J. and Menczer, F., 2009, June. What's in a session: tracking individual behavior on the web. In Proceedings of the 20th ACM conference on Hypertext and hypermedia (pp. 173-182); Shen, Z., Wei, J., Sundaresan, N. and Ma, K.L., 2012, October. Visual analysis of massive web session data. In IEEE symposium on large data analysis and visualization (LDAV) (pp. 65-72). IEEE.

[208] https://developer.mozilla.org/en-US/docs/Web/HTTP/Overview ("Clients and servers communicate by exchanging individual messages (as opposed to a stream of data). The messages sent by the client are called requests and the messages sent by the server as an answer are called responses.").

[209] https://developer.mozilla.org/en-US/docs/Web/HTTP/Overview ("HTTP is a protocol for fetching resources such as HTML documents. It is the foundation of any data exchange on the Web and it is a client-server protocol,

146. The figure below—reproduced from my Opening Class Certification Declaration—plots the timeseries of the transmissions from the web browser to servers when RiteAid's website is loaded. The blue lines in this plot indicate the timing of the various HTTP request/response exchanges between a user's web browser and servers. The red lines in this plot specifically indicate the timing of the HTTP request/response exchanges between a user's web browser and TikTok's server. The plot shows that TikTok Pixel's interception of the contents of communications, such as the full-string URL and Content Information, occur contemporaneously during the loading of the webpage, not after its conclusion.

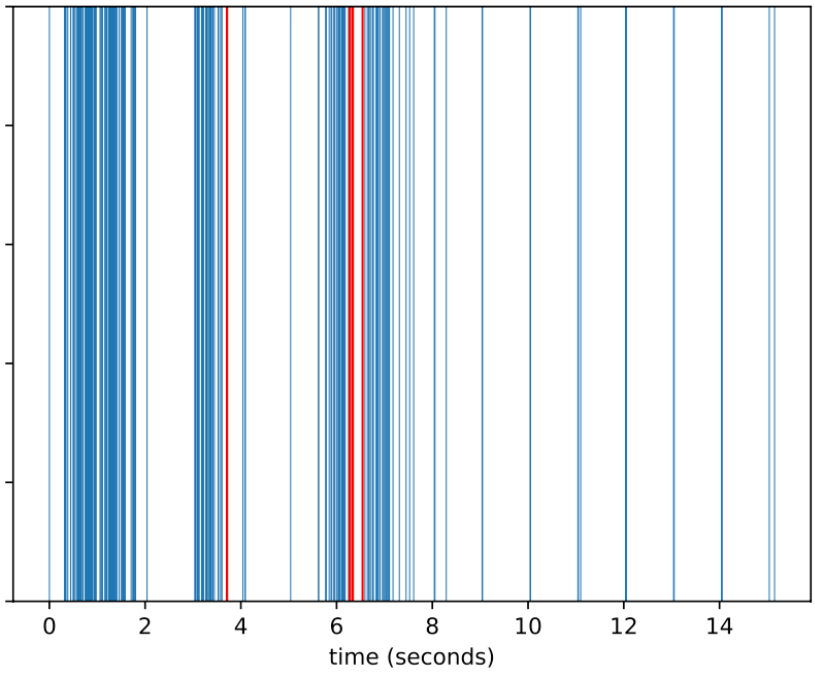


Fig. 6: Timeseries of transmissions when RiteAid's website is loaded
(red lines represent the transmissions from the web browser to TikTok's server that happen contemporaneously during the loading of the webpage)

147. As discussed above in response to the Schnell Reconsideration Report, *see supra* at Sec. III.E, Mr. Schnell's analysis of the timing of interception by Events API continues to ignore

which means requests are initiated by the recipient, usually the Web browser. A complete document is typically constructed from resources such as text content, layout instructions, images, videos, scripts, and more.").

the precise language in TikTok's own documentation that I cited in both my Opening and Reply Class Certification Declarations: "it's **highly recommended to send the event in real-time (without batching)** as soon as it is seen on the advertiser's server" (emphasis in original).[210] TikTok's documentation makes it clear that the intercepted information is sent to TikTok's server in "**real-time**" and "as soon as it is seen on the advertiser's server." TikTok also did not dispute that Defendants ***enable*** advertisers who use the Events API to send events in real-time.[211] Mr. Schnell seems to be conflating the time when TikTok intercepts the content with the time when the intercepted data is received by TikTok's server. Regardless, the time differences between the two is small enough to be described as "**real-time**" by TikTok's own documentation.

V. TIKTOK'S STORAGE AND RETENTION OF NON-TIKTOK USER DATA

148. [redacted]

[210] https://business-api.tiktok.com/portal/docs?rid=p41a33fdhon&id=1771100865818625 (JAE Ex. 70) (filed at Dkt. 266-7) (emphasis in original).

[211] *See* Dkt. 268-3 at Fact 41.

[212] TIKTOK-BG-000439076 at -076.

[redacted]

VI. TIKTOK'S USE OF NON-TIKTOK USER DATA

151. In my Opening Class Certification Declaration and Opening Report, I discussed TikTok's usage of non-TikTok user data collected through the Pixel and Events API.[218]

152. [redacted]

[213] TIKTOK-BG-003050820 at -822; *see also* TIKTOK-BG-000446088 at -088 [redacted]

[214] *See, e.g.*, TIKTOK-BG-000297365, at -365 ("[redacted]

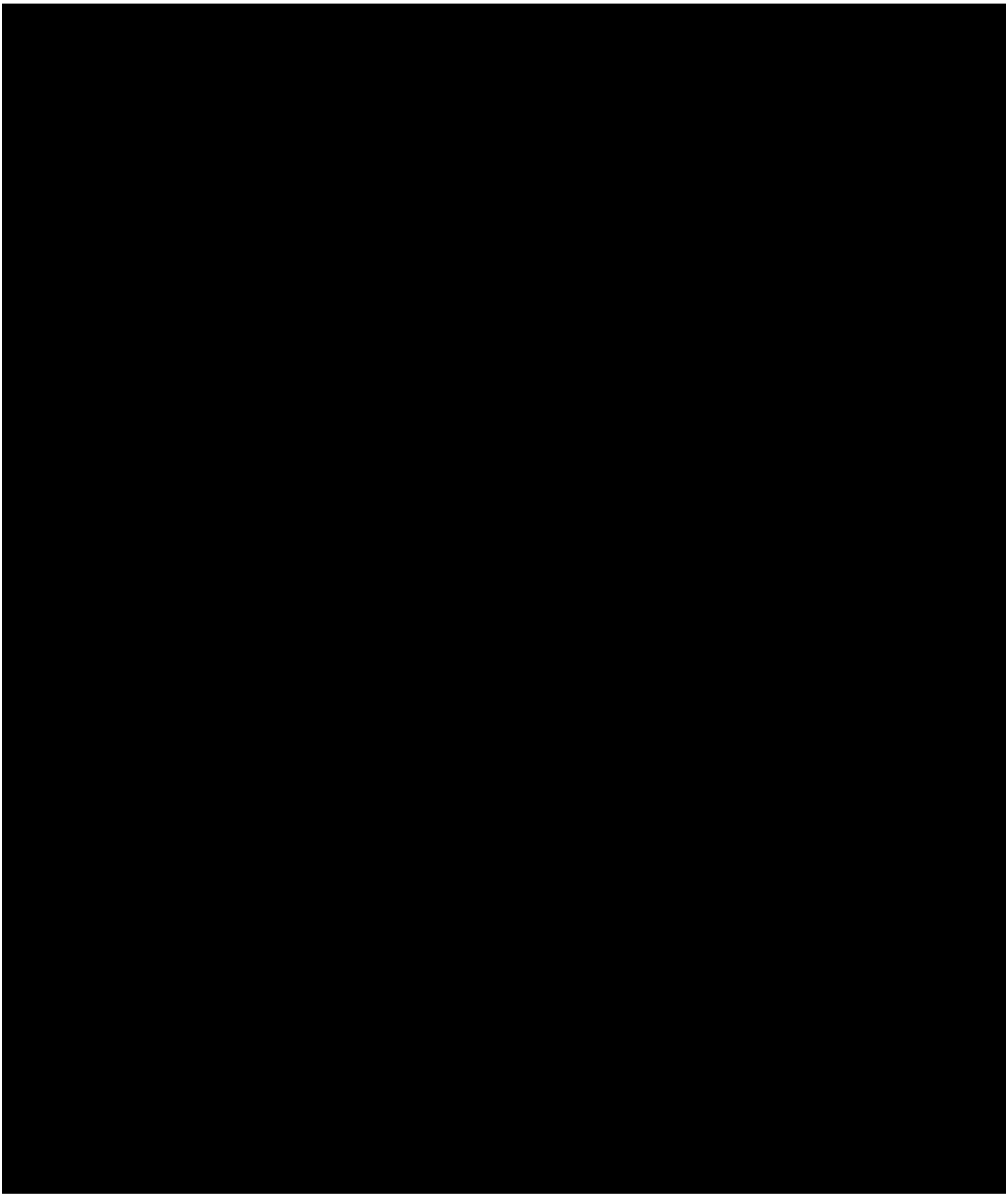
[215] TIKTOK-BG-000447655 at -655 ("[redacted]

[216] TIKTOK-BG-000392758 at -758, 759 [redacted]

see also TIKTOK-BG-000434851 [redacted]

[217] *Id.*

[218] *See* Shafiq Opening Report at ¶¶ 103-108.



[219] TIKTOK-BG-000437740.

[redacted]

154. I understand from counsel that the source code produced by TikTok did not include certain source code repositories that seem to be pertinent to TikTok's use of data collected from TikTok Pixel and Events API, which includes non-TikTok user data. I understand that Plaintiffs have moved for relief from the court for this source code. I reserve the right to amend, modify, and supplement the opinions on TikTok's usage of non-TikTok user data should new information, such as the aforementioned source code, become available to me.

I declare under penalty of perjury under the laws of the United States of America that the foregoing is true and correct. Executed this 11th day of October 2024, in Davis, California.

Zubair

Zubair, Shafiq, Ph.D.

[220] https://www.pangleglobal.com/.

[221] https://ads.tiktok.com/help/article/pangle-placement.

[222] TIKTOK-BG-000694663 at -668. I understand that this document was produced in Chinese, and I have been provided with a certified translation of the document.

[223] TIKTOK-BG-002119116 at -117.